Technical Brief
SWRA030

Understanding and Enhancing Sensitivity in Receivers for Wireless Applications

Edited by Matt Loy *Wireless Communication Business Unit*

Abstract

This technical brief provides an overview of communication receiver sensitivity. One of the most important parameters in determining the overall performance of a communication system, receiver sensitivity translates directly into communication distance and reliability.

A few receiver architectures can offer dependable communication at low cost, if proper design procedures and trade-offs are implemented. RF amplifiers, mixers, and filters are common circuit building blocks for every architecture. System performance is tied to each individual block comprising the receiver. Each circuit generates noise that degrades reception of the desired signal.

Understanding noise sources and the methods of minimizing degradation allows optimal design trade-offs for a given cost. Circuit nonlinearity causes undesired signals to hinder the reception of desired signals. A low-noise system design typically does not produce the best linearity, and high linearity typically produces more noise. A thorough understanding of the receiver RF environment can help your design achieve the proper specifications for optimal noise and linearity.

Contents

Introduction 4
Receiver Architectures 4
    Homodyne (Zero IF) Receiver 4
    Heterodyne Receiver 6
Access Methods and Modulation Schemes 7
    Access Methods 7
    Modulation Schemes 10
Receiver Building Blocks 16
    Antenna 17
    Duplexers 18
    Low-Noise Amplifier (LNA) 19
    RF Filters 19
    Mixers 20
    Local Oscillator 22
    IF Filter 23
    IF Amplifier 23
    Detectors 24
Receiver Sensitivity 24
    Noise of Two-Port Networks 25
    Noise Temperature 29

Noise Figure 32
Cascaded Noise Temperature 37
Cascaded Noise Figure 40
Lossy Two-Port Noise Figure 42
Mixer Noise Figure 44
Minimum Detectable Signal 46
Signal-to-Noise Ratio 48
Sensitivity 50
Phase Noise 50
Receiver Nonlinear Performance 51
Gain Compression 51
Intermodulation Distortion 54
Dynamic Range 60
Blocking 62
Spur-Free Dynamic Range 63
Undesired Spurious Responses 64
Self-Quieters 67
Receiver Design Trade-Offs 67
Summary 70
Glossary 70
References 74

Figures

Figure 1. Homodyne (Zero-IF) Receiver 5
Figure 2. Dual-Conversion Superheterodyne Receiver 6
Figure 3. User Time Slots for TDMA Example 8
Figure 4. Frequency Channels for FDMA Example 8
Figure 5. Modulation Schemes 11
Figure 6. Constellation Diagram (I/Q Plot) 12
Figure 7. OQPSK State Diagram 13
Figure 8. π/4-DQPSK State Diagram 14
Figure 9. 16QAM State Diagram 15
Figure 10. Receiver Components 16
Figure 11. Mixer Driven by LO Signal 21
Figure 12. I/Q Detector 24
Figure 13. Complex Source Connected to Complex Conjugate Load Termination 25
Figure 14. Thermal Noise Visualized as Signal Generators 26
Figure 15. Filter Response Equated to Brick-Wall Filter Response 27
Figure 16. Two-Port Network Connection for Available Gain Definition 29
Figure 17. Available Thermal Noise (Noise Temperature) 30
Figure 18. Noise Output Power Versus Source Temperature (Noise Temperature) 31
Figure 19. Available Thermal Noise (Noise Figure) 32
Figure 20. Noise Output Power Versus Source Temperature (Noise Figure) 33
Figure 21. Noise Figure Quantifies Signal-to-Noise Ratio Degradation 34
Figure 22. Amplifier With 10 dB of Gain and a 5-dB Noise Figure 36
Figure 23. Two-Stage Network (Cascaded Noise Temperature) 38
Figure 24. Cascaded Receiver System (Cascaded Noise Temperature) 39
Figure 25. Two-Stage Network (Cascaded Noise Figure) 40
Figure 26. Cascaded Receiver System (Cascaded Noise Figure) 41
Figure 27. Lossy Two-Port (Noise Figure) 42
Figure 28. Mixer Conversion Noise 44
Figure 29. Single-Sideband Mixer Noise Model 45
Figure 30. Double-Sideband Mixer Noise Model 46
Figure 31. Dual-Conversion Receiver With Associated Noise and Signal Levels 47
Figure 32. Dual-Conversion Receiver and Carrier-to-Noise Ratio 48
Figure 33. Probability of Error Versus Carrier-to-Noise Ratio 49
Figure 34. Double-Sideband Phase Noise Plot of a Signal Source 50
Figure 35. Conversion Limitations Due to Phase Noise 51

Figure 36. Ideal Two-Port P_{out} Versus P_{in} Characteristic 52
Figure 37. Two-Port Gain Versus Input Power Level 53
Figure 38. Receiver Circuits and Intermodulation Distortion 54
Figure 39. Output Power P_{out} for One Fundamental Frequency Signal Versus Input Power P_{in} 56
Figure 40. Fundamental P_o and Third-Order Distortion P_d Output Powers Versus Input Power P_{in} 57
Figure 41. Typical Output Spectrum of a Two-Port Receiver Network 59
Figure 42. Cascaded Receiver Networks 59
Figure 43. Dynamic Range 61
Figure 44. Receiver Blocking 62
Figure 45. Spur-Free Dynamic Range 63
Figure 46. High-Side Injection Example 65
Figure 47. Half-IF Spur 66
Figure 48. Frequency Spectrum for RF and LO Frequencies 66
Figure 49. Receiver Block Diagram 68
Figure 50. Six-Stage Network Combined Into Three-Stage Network 68
Figure 51. New Combined Receiver 69

Table

Table 1. Wireless System Standards 16

Introduction

Optimizing the design of a communications receiver is inherently a process of compromise. Noise limits the smallest signals that a receiver system is capable of processing. The largest signal is limited by distortions arising from nonlinearity of receiver circuits. The smallest and largest signals define the dynamic range of the receiver system. An understanding of the underlying causes of noise enables a receiver designer to get the lowest noise performance receiver during the design process.

Large interfering signals also hinder the reception of small desired signals. The nonlinearities in the receiver networks generate distortion products that fall within the receiver passband. These distortion products prohibit or reduce message reliability. An optimal noise design typically yields less than optimal large signal performance. The best large signal performance suffers from higher noise degradation, which in turn limits the weak signal reception.

For example, a superheterodyne receiver may require filtering between the antenna and front-end mixer to prevent interference or compression from unwanted signals. But a narrowband filter introduces signal losses equivalent to a rise in the receiver noise figure. Adding a low-noise preamplifier to compensate for the filter's losses can lower the noise figure. But the preamplifier also introduces intermodulation (IM) products and raises signal levels to the mixer, possibly degrading mixer spurious performance. However, it is still possible to optimize key receiver performance parameters, such as sensitivity and dynamic range, without compromising other considerations, such as complexity, size, and cost.

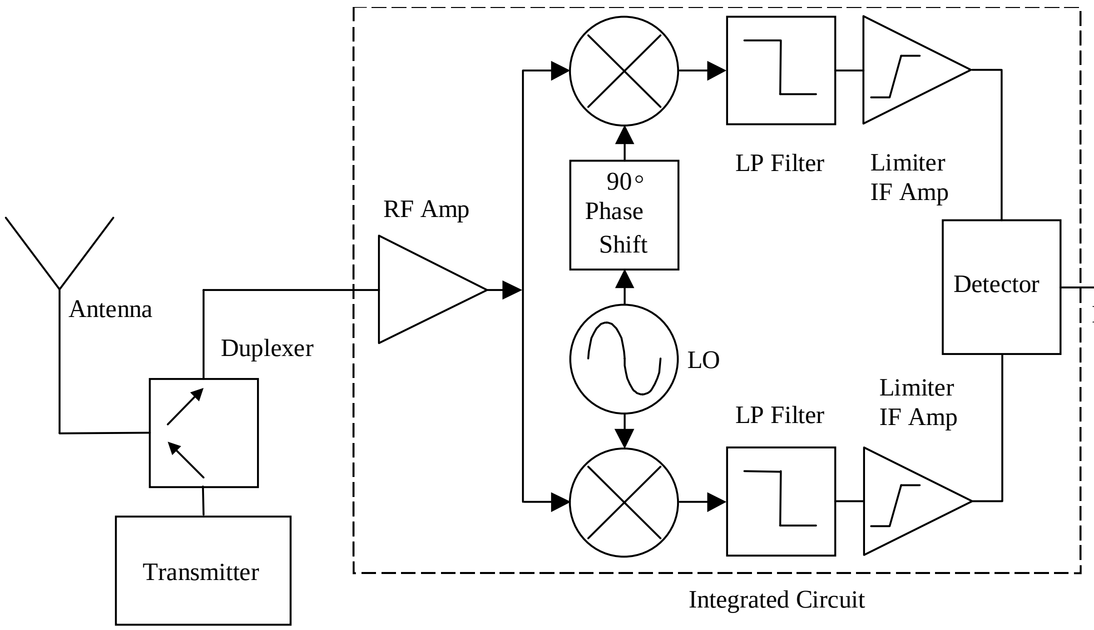
Receiver Architectures

Homodyne (Zero IF) Receiver

A homodyne, direct conversion, or zero-IF (no intermediate frequency) receiver translates the desired RF frequency directly to baseband for information recovery. Baseband is the range of frequencies occupied by the signal before modulation or after demodulation. Baseband signals are typically at frequencies substantially below the carrier frequencies. On the low end of baseband, signals may approach or include DC. The upper frequency limit of baseband depends on the data rate, or speed, at which information is sent, and whether or not special signals called subcarriers are utilized. Figure 1 shows an example block diagram of a zero-IF receiver architecture. A nonlinear circuit known as a mixer translates the RF frequency directly to baseband. A local oscillator (LO) signal, tuned to the same frequency as the desired RF signal, is injected into the mixer. The RF and LO signals mix to produce the baseband frequency. Some digital radios also employ an I/Q mixer to recover baseband information.

When translating directly from RF to baseband, a DC component (along with the band-limited information signal) is realized at the output of the mixer. The DC component (or DC-offset) must be removed to prevent large DC pulses from de-sensitizing the baseband demodulator. The system can either be AC-coupled or incorporate some form of DC notch filtering after the mixer.

Figure 1. Homodyne (Zero-IF) Receiver



The zero-IF receiver can provide narrow baseband filtering with integrated low-pass (LP) filters. Often, the filters are active op-amp-based filters known as *gyrators*. The gyrators provide protection from most undesired signals. The gyrator filters eliminate the need for expensive crystal and ceramic IF filters, which take more space on a printed circuit board. The zero-IF topology offers the only fully integrated receiver currently possible. This fully integrated receiver solution minimizes required board real estate, the number of required parts, receiver complexity, and cost. Most zero-IF receiver architectures also do not require image reject filters, thus reducing cost, size, and weight.

Zero-IF receiver limitations require tighter frequency centering of the LO and RF frequencies. Significant offsets in the RF or LO frequencies degrade bit error rate. When the desired signal is above the VHF range, the zero-IF design becomes more complex, partially due to these frequency offset problems. One solution for higher frequency zero-IF designs is to add automatic frequency control (AFC). AFC prevents the centering problem by adjusting the frequency of the LO automatically.

Performance is typically limited in a zero-IF architecture in several ways. Sensitivity and rejection to some undesired signals, such as intermodulation distortion, can be difficult to achieve. The active gyrator filters compress with some large undesired signals. Once the gyrator is compressed, filter rejection is reduced, thus limiting protection. Zero-IF receivers typically require an automatic gain control (AGC) circuit to protect against large signal interference that compresses the gyrator filters.

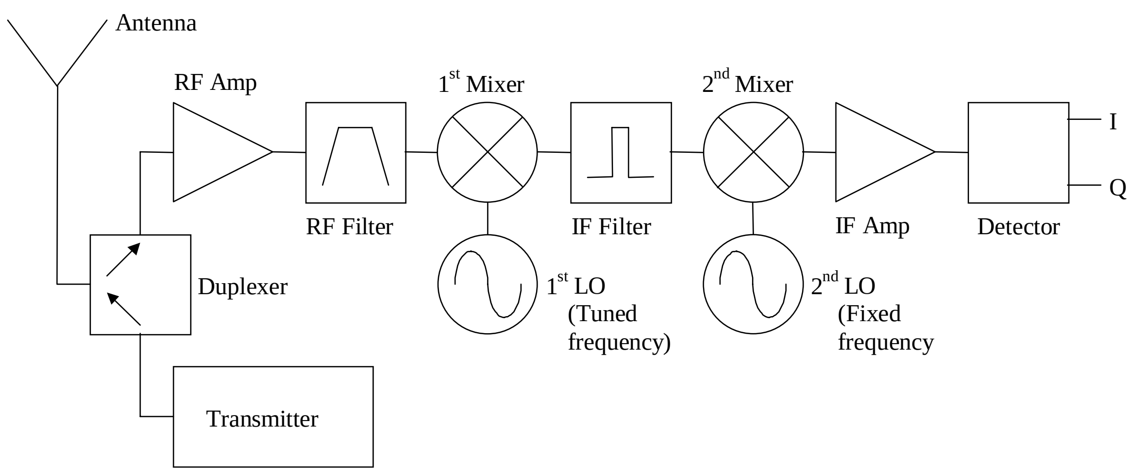
Because the local oscillator is tuned to the RF frequency, self-reception may also be an issue. Self-reception can be reduced by running the LO at twice the RF frequency and then dividing-by-two before injecting into the mixer. Because the Zero-IF local oscillator is tuned to RF frequencies, the receiver LO may also interfere with other nearby receivers tuned to the same frequency. The RF amplifier reverse isolation, however, prevents most LO leakage to the receiver antenna.

Heterodyne Receiver

A heterodyne receiver translates the desired RF frequency to one or more intermediate frequencies before demodulation. Modulation information is recovered from the last IF frequency. Figure 2 shows a dual-conversion superheterodyne receiver. Mixers translate the RF signal to IF frequencies. LO signals tuned at a particular spacing above or below the RF signal are injected into the mixer circuits. The RF and LO signals mix to produce a difference frequency known as the IF frequency. The result is a dual-conversion receiver, described as such because of the two down-conversion mixers.

The advantages that a heterodyne receiver has over a zero-IF receiver include better immunity from interfering signals and better selectivity.

Figure 2. Dual-Conversion Superheterodyne Receiver



Narrow-bandwidth passive IF filtering is typically accomplished using crystal, ceramic, or SAW filters. These filters offer better protection than the zero-IF receiver gyrator filters against signals close to the desired signal because passive filters are not degraded by the compression resulting from large signals. The active gyrator circuit does not provide such protection. However, the price for improved protection is larger physical size and required printed circuit board real estate.

Undesired signals that cause a response at the IF frequency in addition to the desired signal are known as *spurious responses*. Spurious responses must be filtered out before reaching mixer stages in the heterodyne receiver. One spurious response is known as an *image frequency*. An RF filter (known as a preselector filter) is required for protection against the image unless an image-reject mixer is used. An advantage of the zero-IF receiver is that no image exists and an image-reject filter (or image-reject mixer) is not required.

Additional crystal-stabilized oscillators are required for the heterodyne receiver. Superheterodyne receivers typically cost more than zero-IF receivers due to the additional oscillators and passive filters. These items also require extra receiver housing space. However, a superheterodyne receiver's superior selectivity may justify the greater cost and size in many applications.

Access Methods and Modulation Schemes

A single-duplex system allows transmission and reception in one direction at one time. A full-duplex system allows simultaneous transmission and reception in both directions. In wireless systems, multiple-access techniques allow parallel transmissions in both directions.

Multiple-access methods are based on any of four domains:

- Spatial
- Frequency
- Time
- Code

In time division multiple access, different signals are multiplexed in non-overlapping time slots and are differentiated by their times of arrival at the receiver.

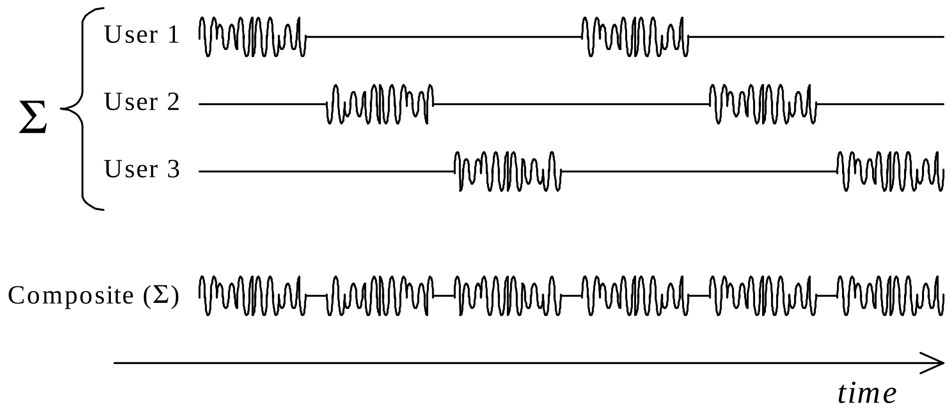
In the code domain, signals have little cross-correlation even though they might be transmitted at the same time and within the same frequency channel. Correlators are used to separate the uniquely coded signals. The access scheme is known as code-division multiple access.

Access Methods

TDMA

In time division multiple access (TDMA), time slots differentiate users. The available radio spectrum is divided into time slots. Users can either transmit or receive in their dedicated time slot. Time slots for *N* number of users are collected into a periodic frame, with *N* time slots per frame. Because TDMA data is transmitted in bursts, transmission for a given user is not continuous. Temporal synchronization between a TDMA transmitter and receiver using time gating permits reception of a specific user's time-slot data, essentially turning the receiver on and off at the appropriate times. The users essentially "take turns" using the same radio channel. Figure 3 shows an example TDMA system with three users. The composite signal indicates that the channel is in use for the complete length of time.

Figure 3. User Time Slots for TDMA Example

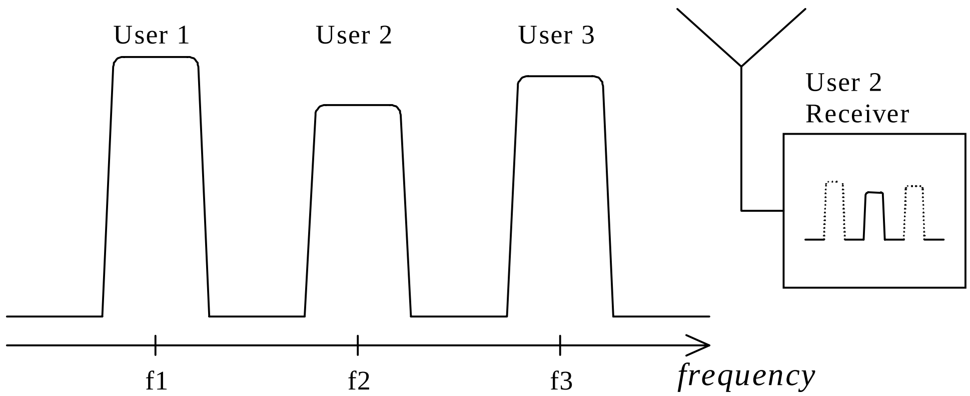


TDMA is sometimes combined with time division duplex (TDD) or frequency division duplex (FDD). With TDD, half of the frame's time slots are used for transmission and half are used for reception. With FDD, different carrier frequencies are used for TDMA transmission and reception, even though similar time slots may be used for both functions.

TDMA systems using FDD usually add several time slots of delay to separate a particular user's transmission and reception time slots and to synchronize transmitters and receivers. The guard time between transmissions can be designed as small as the system's synchronization permits. (Guard times of 30 to 50µs between time slots are common in practical TDMA systems). All mobile units must be synchronized to the base station to within a fraction of the guard time.

Narrowband TDMA is used in the Global System for Mobile Communications (GSM). The IS-136 TDMA standard employs 30-kHz FDMA (frequency domain multiple access) channels subdivided into three time slots (six time slots for ½ rate) with one time slot required for each call made.

Figure 4. Frequency Channels for FDMA Example



Users of FDMA systems divide the frequency spectrum into narrow bandwidth channels where each user is assigned a specific frequency or frequencies for transmission and reception. The receiver demodulates information from the desired channel and rejects other signals nearby. Figure 4 shows three users of an allocated frequency band where the receiver for User 2 is shown. The channel, f2, is demodulated while rejecting channels f1 and f3.

Efficiency in a TDMA system is measured by the percentage of transmitted data containing information versus overhead. The frame efficiency is the percentage of bits per frame that contain transmitted data. The transmitted data may contain source and channel coding bits, so that the end-user efficiency of the system is usually less than the frame efficiency.

CDMA

Code-division multiple access (CDMA) systems are either direct-sequence spread spectrum (DSSS), which use orthogonal or un-correlated pseudorandom-noise (PN) codes to differentiate signals that overlap in both frequency and time, or frequency-hopping spread spectrum (FHSS), in which signals are randomly hopped about different portions of an available spectrum.

In DSSS CDMA, a narrowband message signal is multiplied by a very large-bandwidth PN spreading signal with a chip rate that is orders of magnitude larger than the data rate of the message signal. (The chip period is the inverse of the spreading rate.) A large spreading rate can minimize the effects of multipath distortion, such as channel signal fading.

Each user has a unique pseudorandom code. Due to the orthogonality of the codes, all other codes appear as noise to a given user. A matched filter extracts a specific user's signal, with little or no output signals resulting from other users. Armed with the proper code for that user, a CDMA receiver can effectively extract a user’s signals from a channel with multiple signals at the same relative amplitude level.

For CDMA systems to work effectively, the power levels of mobile units seen by the base-station receiver must be approximately equal. CDMA achieves this balance using dynamic power control techniques. The total power of multiple users at a receiver determines a CDMA system’s noise floor after decorrelation. Unless the power of each user within a cell is controlled so that they are approximately equal at the base station receiver, the strongest received mobile signal can dominate the base-station’s demodulator.

Signals rising above the level of the other signals increase the noise level at the base station receiver. Higher-level signals decrease the probability that a desired signal will be received and decorrelated. As a result, CDMA systems employ power control at each base station to control the signal level received from each mobile unit, making them all approximately equal. In this way, a mobile unit close to the base station will not overpower the base station for a user much further away.

Power control is implemented by rapidly sampling the received signal strength indicator (RSSI) level from each mobile unit and then sending a power change command over the forward radio link to each unit. Unfortunately, out-of-cell mobile units can still provide interference for the base station.

On the up-link (from a mobile unit to the base station), power control must be accurate to within 1 dB and fast enough to compensate for channel fading. Given a vehicle traveling at 55 mph, the Doppler bandwidth or rate is about 100 Hz (the channel changes characteristics 100 times each second). This requires power-control bits of about 10b/Hz or about 1kb/s of control data. The power-control data must accommodate a dynamic range of about 80 dB.

To its credit, CDMA enjoys a soft capacity limit; that is, an increase in the number of subscribers raises the noise floor in a linear manner. As a result, the number of users in a CDMA system is not absolutely limited. The system performance gradually degrades as the number of users increases and improves as the number of users decreases. As noted, large spreading rates can negate the effects of multipath distortion. RAKE receivers are often used in CDMA systems to improve reception by collecting several time-delayed versions of the desired signal.

On the negative side, self-jamming can be a problem in a CDMA system because the spreading sequences of different users are not exactly orthogonal, causing despreading for a particular PN code. In addition, power control can be difficult to implement when undesired subscribers (such as from adjacent cells) are not rejected, and their power levels add to the system's total noise floor.

CDMA cellular or PCS system capacity is determined by a number of factors, including the processing gain, the effective bit-energy-to-noise ratio (E_b/N_o), frequency reuse, and the number of sectors in a cell. Because many subscribers operate within the same 1.25-MHz-wide CDMA channel, frequency reuse efficiency is increased by improving the signal-to-noise ratio (SNR).

Modulation Schemes

Sending low frequency information without using wires requires a transmitter and receiver. If the information were sent directly without wires through space, the receiver and transmitter antennas would have to be enormously large. Antenna sizes are reduced as frequency increases. A high frequency signal is therefore encoded with the information to be sent. The unmodulated RF carrier is shown in Figure 5a. The RF carrier has three available characteristics for information encoding:

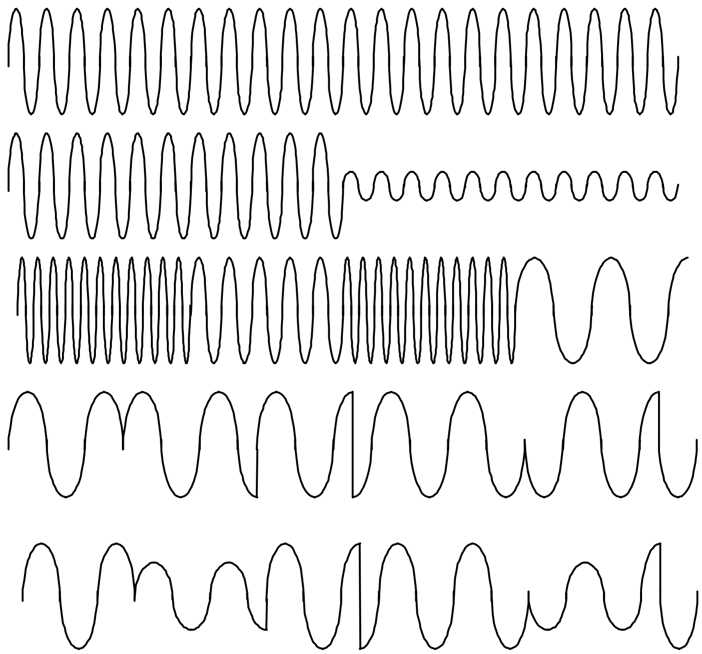
- Amplitude
- Frequency
- Phase

Figure 5b shows a change in the signal amplitude known as *amplitude modulation*. In amplitude modulation, the carrier amplitude is varied instantaneously with the information. The AM radio broadcast bands, covering 550 kHz to 1600 kHz, use this method.

A change in the carrier frequency known as *frequency modulation* is shown in Figure 5c. The carrier amplitude is held constant for frequency modulated signals. The FM radio broadcast band, from 88 MHz to 108 MHz, utilizes frequency modulation, varying the carrier frequency proportionately with the information being sent. Some digital radios, such as paging services, use FM modulation to send information. Here the modulation format is known as frequency shift keying (FSK). In FSK, the frequency shifts between two or more values for a particular bit or symbol of information.

Figure 5d shows a change in the carrier phase known as *phase modulation*. Digital radios, such as CDMA cellular, use phase modulation techniques such as quadrature phase-shift keying (QPSK).

Figure 5. Modulation Schemes



An I/Q modulator allows any parameter, amplitude, frequency, or phase to be simultaneously manipulated, for example, both amplitude and phase as shown in Figure 5e. This simultaneous manipulation allows each transmitted symbol to represent more than one bit of data. The signal's amplitude and phase is modulated by separating the signal into a set of independent channels called I and Q components. To generate the I and Q components, the signal is fed into two separate mixers driven by the same local oscillator. The local oscillator for one of the components is phase-shifted by 90 degrees, producing an in-phase (I) component and a quadrature-phase (Q or 90°) component. An I/Q modulator is also used to generate constant-envelope phase modulation, such as GMSK, or phase modulation, such as QPSK. Other modulation schemes, such as 16QAM, employ both amplitude and phase modulation provided by I/Q modulation techniques.

Sending information by modulating the carrier amplitude (AM) or the frequency (FM) requires simple hardware but is inefficient because the frequency spectrum limits the number of users. More complicated digital modulation methods, such as 16QAM, increase hardware complexity but use the limited frequency spectrum much more efficiently, thus allowing more users in a given bandwidth. Because digital modulation schemes offer many benefits and capacity improvements as compared to analog modulation, they are used in the majority of current wireless standards.

The spectral efficiency of a digital radio depends on the modulation scheme. The signal bandwidth for a digital communications channel depends on the symbol rate as opposed to the bit rate. Symbol rate is the ratio of the bit rate to the number of bits transmitted with each symbol, or in equation form:

(1) $$\text{Symbol Rate} = \frac{\text{Bit Rate}}{\text{\# of Bits per Symbol}}$$

Frequency Modulation (FM)

Frequency modulation varies the frequency of a constant amplitude carrier as a function of the information amplitude. Analog systems typically employ FM to send information. Early cellular telephone systems transmit on one frequency band and receive on another band, thereby allowing a full-duplex communication system. The Advanced Mobile Phone System (AMPS) handset receives an FM signal on a channel in the 869-MHz to 894-MHz frequency band and transmits on a channel from 824 MHz to 849 MHz. The AMPS network utilizes the frequency spectrum inefficiently, which limits the number of subscribers.

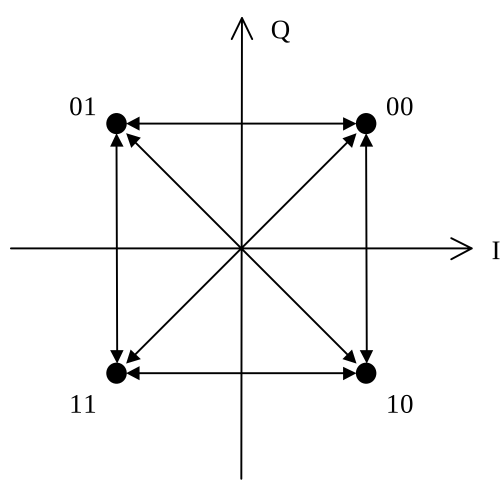
QPSK/OQPSK

Quadrature phase-shift keying (QPSK) and offset quadrature phase-shift keying (OQPSK) are digital modulation forms requiring an I/Q modulator in the transmitter and an I/Q demodulator in the receiver. CDMA systems utilize QPSK and OQPSK modulation methods. By transmitting two bits per symbol, QPSK is more efficient than binary phase-shift keying (BPSK), which transmits one bit per symbol.

Allowing a second orthogonal channel over a basic BPSK-modulated signal improves bandwidth efficiency. Placing a second BPSK signal in quadrature with the first without causing interference to either of the two signals effectively doubles the bandwidth efficiency. This process is known as *QPSK modulation*.

The data stream travels at a bit rate of 1/T b/s and is separated into two data streams at the modulator input port. The separated data streams are even and odd bits. The four states are +45°, +135°, -45°, and -135°. On an I/Q plot, this modulation form appears as four equally spaced points separated by 90°. Each of the four phase states represents two data bits. Figure 6 shows a state diagram, called a constellation diagram (or I/Q plot), for QPSK.

Figure 6. Constellation Diagram (I/Q Plot)



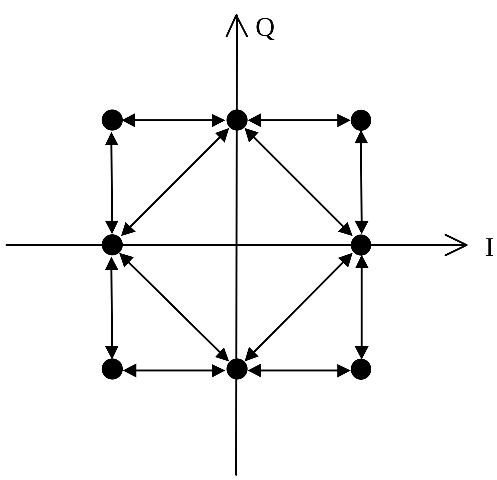
The symbol rate is the bit rate divided by the number of bits per symbol, given by equation (1). The symbol rate for QPSK is one half the bit rate. Twice as much information can be carried with QPSK modulation as compared to BPSK modulation, which has a symbol rate equal to the bit rate. Compared to BPSK, the required channel bandwidth in QPSK for a data rate of 24 kb/s is nearly 100 kHz—an improvement by a factor of two in bandwidth efficiency.

The carrier phase in QPSK can change only once every 2T periods because the carrier phase over that interval can be only one of the four possible phase states. In the next interval, if neither bit stream changes sign, the carrier phase remains the same. If one of the phase components changes, a 90° phase shift will occur. A shift in both phase components results in a phase shift of 180°. A shift again in both phase components leads to a 180° phase reversal.

Delaying the odd-bit data stream by a one-half-bit interval with respect to the even bit produces OQPSK modulation. This offset reduces the range of phase transitions to 0° and 90°, with the transitions to those phase states per 2T interval occurring twice as often as in QPSK. This offset also reduces the envelope fluctuation at the modulator output. Figure 7 shows an OQPSK constellation diagram.

Because of fast envelope fluctuations in QPSK, the high-frequency components removed during filtering (for adjacent-channel interference) in a cellular system can be regenerated by a nonlinear amplifier. Because of the absence of fast phase transitions in OQPSK, a nonlinear amplifier will regenerate less high-frequency spurious components. Thus, for cellular systems, the amplifier in the handset can operate in a more efficient, less linear region and increase talk time.

Figure 7.   OQPSK State Diagram



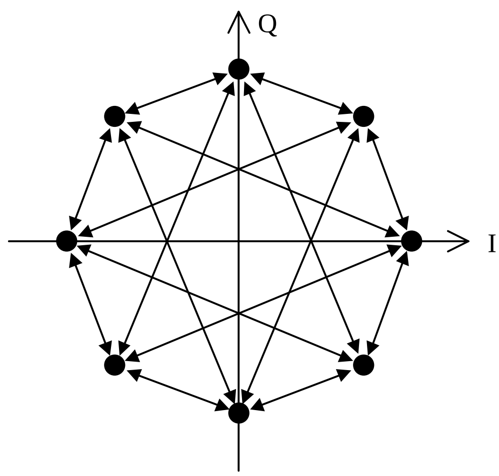
π/4-DQPSK

A compromise to standard QPSK is π/4-DQPSK (differential QPSK) modulation. Its phase fluctuations are restricted to ±π/4 and ±3π/4. This restriction is worse than ±π/2 for OQPSK but allows noncoherent detection because of the different phase states. The spectral efficiency of π/4-DQPSK modulation is about 20 percent better than that of GMSK.

The π/4-DQPSK modulation is essentially π/4-shifted QPSK with differential encoding of symbol phases. Differential coding minimizes the loss of data due to phase slips but also results in the loss of a pair of symbols when channel errors occur. This effect can account for about a 3-dB loss in E_b/N_o relative to coherent π/4-QPSK.

A π/4-DQPSK signal constellation comprises symbols corresponding to eight phases (0 to 360° every 45°). These eight phase points can be formed by superimposing two QPSK signal constellations offset by 45° relative to each other, as shown in Figure 8. During each symbol period, a phase angle from one of the QPSK constellations is transmitted. The two constellations are used alternately to transmit every pair of bits. As a result, successive symbols have a relative phase difference that is one of four phases. This process can be performed by differential encoding of the source bits and then mapping them into absolute phase angles, or by directly mapping the pairs of input bits into relative phases. IF signals with π/4-DQPSK modulation can be readily demodulated with a straightforward I/Q architecture. Again, because of the modulation envelope fluctuation, the handset amplifier must be operated in a linear region to reduce spectral regrowth.

Figure 8. π/4-DQPSK State Diagram



GMSK

Minimum-shift-keying (MSK) modulation is a form of OQPSK in which the symbol pulse shape is a half-cycle sinusoid rather than the usual rectangular pulse. Simply described, MSK is OQPSK with sinusoidal pulses. With its constant-envelope property, the power amplifier may work at saturation without significantly distorting the modulated signals. This allows the use of higher efficiency power amplifiers in the system transmitters.

MSK has a constant envelope and is phase continuous while QPSK is phase discontinuous. For each bit interval, the carrier has a phase shift as a function of time of ±π/2. Values of +1 and –1 correspond to higher and lower frequencies, respectively, in the modulation scheme. The carrier frequency is never actually transmitted; only these slightly higher and slightly lower modulated forms of the carrier frequency are transmitted.

The Gaussian minimum-shift keying (GMSK) modulation scheme reduces signal sidelobe power level and mainlobe width by filtering the rectangular pulses before modulation using a baseband Gaussian-shaped filter. This controlled amount of inter-symbol-interference (ISI) further reduces the carrier's phase discontinuities. This approach can satisfy an adjacent-channel interference requirement of 60 dB or more. (The ratio of spurious content in the adjacent channel to that of the desired channel is 60 dB). This constant-envelope modulation form suffers minimal spectral spreading due to the effects of nonlinear amplification, making it suitable for systems with difficult adjacent-channel interference requirements.

16QAM Quadrature Amplitude Modulation

16QAM, commonly used in Motorola iDEN radio networks and microwave digital radios, offers four values for I and four values for Q, yielding 16 possible states, as shown in Figure 9. 16QAM sends four bits per symbol. The signal can transition from any state to any other state. 16QAM is more spectrally efficient than BPSK, QPSK, OQPSK, and π/4-DQPSK.

This approach encounters increased problems with inter-symbol interference because the nonlinearity of the transmitter and receiver paths along with noise may cause one symbol to be interpreted as another symbol, thus causing an error.

Figure 9. 16QAM State Diagram

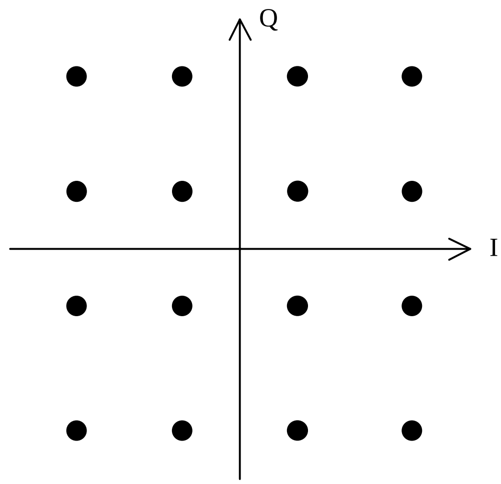


Table 1 summarizes wireless system standards.

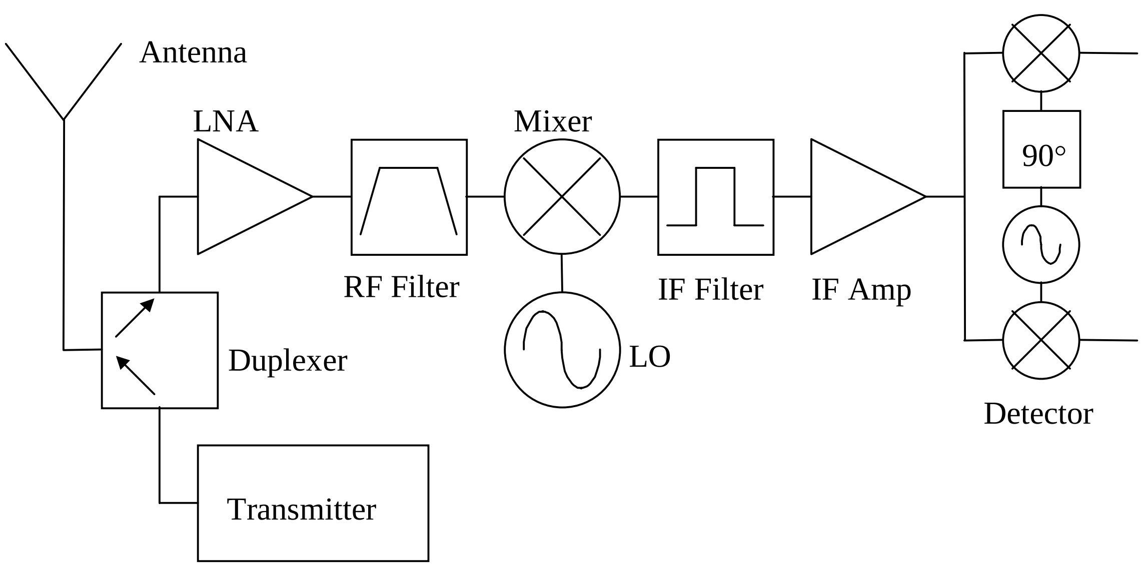
Table 1. Wireless System Standards

Parameter	E-GSM900	DCS1800	IS-136	IS-95	PDC
Tx (MHz)	880-915	1710-1785	824-849 1850-1910	824-849	940-956, 1477-1501
Rx (MHz)	925-960	1805-1880	869-894 1930-1990	869-894	810-826, 1429-1453
Access method	TDMA	TDMA	TDMA	CDMA	TDMA
Modulation	GMSK	GMSK	π/4-DQPSK	QPSK BPSK	π/4-DQPSK
Carrier spacing (kHz)	200	200	30	1250	25
Duplex	FDD	FDD	FDD	FDD	FDD

Receiver Building Blocks

Optimal receiver sensitivity begins by properly choosing receiver components. A receiver system consists of numerous active and passive function blocks or components, each of which contributes to the system's overall signal gain/loss and noise figure. These components include antennas, amplifiers, filters, mixers, and signal sources, as shown in Figure 10.

Figure 10. Receiver Components



Antenna

The antenna provides an interface between free space and the receiver input. An antenna is sometimes considered a matching circuit that matches the free-space impedance to the receiver input. Antennas are characterized by a number of performance parameters, such as bandwidth, gain, radiation efficiency, beamwidth, beam efficiency, radiation loss, and resistive loss.

Antenna gain is often defined relative to a theoretical ideal isotropic antenna that radiates energy equally in all directions. With this convention, an antenna rated for 3-dBi gain in one direction has 3 dB more gain than an isotropic antenna in that same direction. An antenna is characterized relative to a half-wave dipole antenna having a gain of 2 dBi. A gain value of 3 dBd indicates a signal level 3 dB above that achieved by a dipole antenna (5 dBi).

An antenna's gain is defined as either the power gain or directive gain. The *directive gain* is the ratio of the radiation intensity in a given direction to the radiation intensity of an isotropic radiator. *Directivity* is the value of the directive gain in the direction of maximum radiation. Radiation efficiency is given by equation (2).

(2) $$E = \frac{P_r}{P_t}$$

where:

P_r is antenna-radiated power.
P_t is the total power supplied to the antenna.

Antenna beamwidth is the angle formed by the radiating field where the electric field strength or radiation pattern is 3 dB less than its maximum value. The *antenna main lobe* is the energy traveling in the primary direction of propagation. Energy traveling to the sides of the primary propagation direction is called *side lobe energy*. Energy flowing to the rear of the antenna is called the *back lobe radiation*.

Effective radiated power (ERP) is equal to the input power to an antenna P_i multiplied by the transmitter antenna gain (G_t) with respect to a half-wave dipole.

(3) $$ERP = P_i \times G_t$$

The effective isotropic radiated power (EIRP) is the actual power delivered to the antenna multiplied by its numeric gain with respect to an isotropic radiator. Because the gain for an isotropic antenna differs from the gain for a dipole antenna (2 dB), EIRP is about 2 dB higher than the ERP for a given transmission system.

When an antenna is used with a receiver, its noise temperature provides a useful measure of its noise contributions to the system, notably in satellite systems. The noise power available at the output feed port in watts is given by equation (4).

(4) $$P_n = KT_aB$$

where

$K \equiv$ Boltzmann's constant = 1.38×10^{-23} J/°K.
$T_a \equiv$ antenna noise temperature in degrees °K.
$B \equiv$ noise equivalent bandwidth in Hertz (Hz).

The antenna matching circuit matches the antenna output impedance to the receiver input. A gain versus noise figure trade-off is selected for optimum receiver sensitivity when the antenna is connected directly to an amplifier input. When connecting to a filter or diplexer circuit, optimum sensitivity is obtained with a power match. The antenna matching components should have the lowest loss possible because loss increases noise figure as detailed in the section, *Noise Figure* .

Proximity effects such as metal structures, human bodies, and other nearby objects alter antenna performance. Antenna gain, impedance, radiation pattern, and resistive loss are all affected by proximity effects.

Duplexers

A duplexer allows simultaneous transmitter and receiver operation with a single antenna. The ideal duplexer perfectly isolates the receiver and transmitter from each other while providing lossless connection to the antenna for both.

Duplexers are constructed in a variety of ways. One method uses an RF switch to toggle the antenna back and forth between the receiver and transmitter. The toggling occurs with a “push-to-talk” (PTT) button or is time-shared automatically by a microprocessor. For this method, a separate preselect filter is often needed on the receive path prior to the LNA to limit the incoming frequencies to the band of interest.

Another method uses two singly terminated designed filters known as a diplexer. Singly terminated filters allow either a low impedance or high impedance termination on one port. The two filters are connected at this port, thus forming a three-terminal network. Diplexers are employed where the transmitter and receiver frequencies are in different bands. The passband of one filter is the stopband for the other. For example, one filter may be a lowpass whereas the other may be a highpass. The receiver-band filter rejects transmitter-band signals and the transmit-band filter rejects receiver-band signals. For this method, the diplexer can also act as the preselect filter.

Circulators are also used as duplexers. A circulator is a three-port device that allows RF signals to travel in only one direction. Thus, a signal connected to port 1 transfers to port 2 unimpeded and is isolated from port 3. A signal incident on port 2 transmits to port 3 and is isolated from port 1. A signal input at port 3 travels to port 1 while being isolated from port 2.

The ideal circulator provides 0-dB insertion loss from port 1 to port 2; port 2 to port 3; and port 3 to port 1. The ideal circulator also provides infinite isolation between port 1 to port 3; port 3 to port 2; and port 2 to port 1. High isolation between the receiver and transmitter minimizes leakage of transmitter power into the receiver and prevents receiver signals from entering the transmitter.

In addition, a duplexer is characterized by insertion loss. Excessive insertion loss degrades the noise figure of the system, which in turn inhibits signal-to-noise ratio performance.

Low-Noise Amplifier (LNA)

An RF amplifier is a network that increases the amplitude of weak signals, thereby allowing further processing by the receiver. Receiver amplification is distributed between RF and IF stages throughout the system. The ideal amplifier increases the amplitude of the desired signal without adding distortion or noise. Numeric gain factor is defined as the output signal power S_{out} divided by the input signal power S_{in}. Numeric gain factor G is defined mathematically by equation (5).

(5) $$G = \frac{S_{out}}{S_{in}}$$

An increase in signal amplitude indicates a numeric gain factor greater than unity. Conversely, a decrease in signal amplitude indicates a numeric gain factor less than unity. A gain factor of unity indicates no change in signal amplitude processed by the two-port network. The RF amplifier provides a gain factor greater than unity. (Gain factor is also used to describe losses in the system.) Numeric gain factor is converted to gain in dB (G_{dB}) using equation (6).

(6) $$G_{dB} = 10\log(G) = 10\log\left(\frac{S_{out}}{S_{in}}\right)$$

Amplifiers unfortunately add noise and distortion to the desired signal. The first amplifier after the antenna in a receiver chain contributes most significantly to system noise figure, assuming low losses in front of the amplifier. Adding gain in front of noisy networks reduces noise contribution to the system from those networks. Subsequent stages have less and less influence on the overall noise figure of the system. Increasing gain from the low-noise amplifier improves system noise figure.

On the other hand, too much gain compresses circuits further back in the receiver chain. The receiver design must make a trade-off between system noise figure and gain.

A low-noise amplifier is typically constructed from active devices operated in the "linear range". The active device is operated in its "linear range," but the output signal isn't perfectly linear. Thus, distortion is added to the amplified signal by nonlinearities of the transistor. Gain compression, harmonic distortion, cross-modulation, and intermodulation distortion directly result from amplifier nonlinearity.

RF Filters

An RF filter is a network that allows a range of RF frequencies to pass. The wanted frequency range is known as the passband. The filter blocks RF signals outside of the passband. The blocked area is known as the stopband. The perfect RF filter passes desired RF signals unimpeded and infinitely attenuates signals in the stopband. Many receivers use two RF filters in the receive path: a preselect filter (before the LNA) and an image-reject filter (after the LNA). The preselect filter prevents signals far outside of the desired passband from saturating the front end and producing intermodulation distortion products related to those signals at far away frequencies only. The image-reject filter rejects signals such as the first image, half-IF, and local oscillator spurious responses. However, the RF filters typically provide little protection against third-order intermodulation distortion produced by "close-in" signals. The nonideal RF filter degrades receiver noise figure by adding loss to the desired signal.

Filters are designed with one of several responses, such as Butterworth, Chebyshev, elliptic, and Bessel responses. A Butterworth response has flat gain across the passband (maximally flat filter) with a gradual roll-off in the transition region from the passband to the stopband. Its phase response is nonlinear about the cutoff region with a group delay that increases slightly toward the band edges.

Chebyshev filters have equal-ripple response in their passbands with better selectivity than a Butterworth for the same order filter but worse phase response because of group-delay variations at the band edges.

An elliptic filter achieves the maximum amount of roll-off possible for a given order but with an extremely nonlinear phase. This type of filter exhibits equal-ripple amplitude response in both passband and stopbands. Its group delay is lower than that of a Chebyshev filter but varies rapidly at the band edges.

A Bessel response is maximally flat in phase within the passband, with a relative equal-ripple response but less than ideal sharpness in the roll-off region.

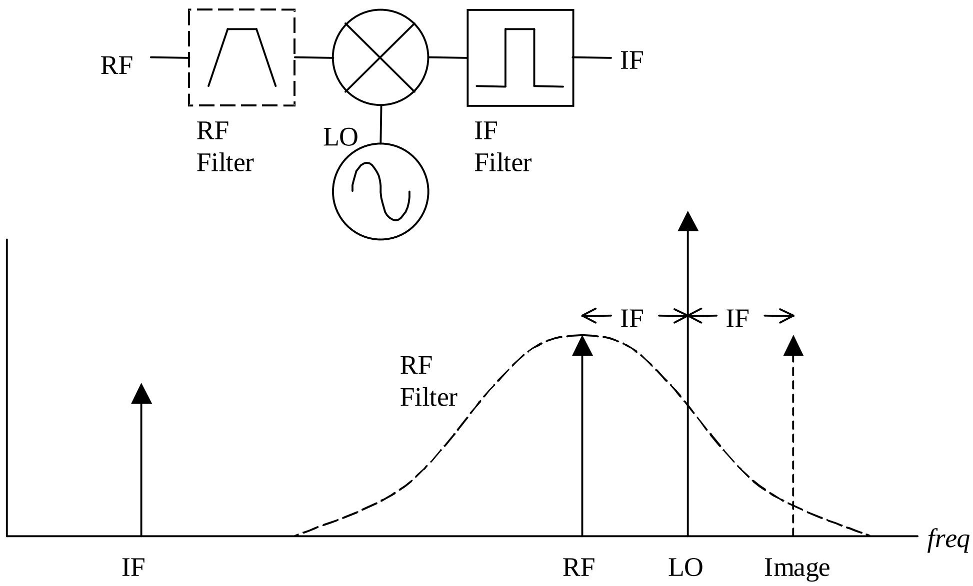
Preselect filters are often ceramic, lumped-element LC or SAW. Ceramic filters typically offer lower loss and are less expensive, but their larger size is a drawback. Image-reject filters are usually either ceramic or SAW devices in cellular receiver front ends. SAW filters are similar to digital FIR filters and have sharp roll-off and linear phase.

Mixers

Mixer circuits translate an RF frequency to both a higher and lower intermediate frequency (IF) value. Figure 11 shows a mixer driven by an LO signal. An IF filter selects either the higher or lower (sum or difference) output frequency. One frequency is passed while the other is rejected. Selecting the higher frequency is up-conversion; selecting the lower frequency is down-conversion. Down-conversion is shown in Figure 11. The translation uses a local oscillator (LO) signal that mixes with the RF frequency. The RF and LO frequencies are spaced apart by an amount equal to the IF frequency. High-side injection has the LO frequency above the RF frequency. Low-side injection puts the LO below the RF frequency. High-side injection is shown in Figure 11.

The down-conversion signal is the frequency difference between the RF and LO frequencies. An additional signal called the *image frequency* has the opportunity to mix with the LO for an IF response. Image-reject RF filters (by filtering the image prior to the mixer) and image-reject mixers (by reducing the image component during the mixing process) provide protection against the image. The image and LO frequencies are spaced apart by an amount equal to the IF frequency. The difference between the LO and image frequencies produces an IF response coincident with the IF response due to the RF signal. An RF (preselector or image) filter in front of the mixer rejects an image frequency that interferes with the desired RF signal.

Figure 11. Mixer Driven by LO Signal



The mixer design uses nonlinear devices, such as diodes or transistors. Using diodes, the mixer is passive and has a conversion loss. Using active devices, such as transistors, a conversion gain is possible. A variety of circuit topologies exist for mixers. A single-ended mixer is usually based on a single Schottky diode or transistor. A balanced mixer typically incorporates two or more Schottky diodes or a Schottky quad (four diodes in a ring configuration). A balanced mixer offers advantages in third-order intermodulation distortion performance compared to a single-ended mixer because of the balanced configuration.

While the mixer operates within its linear range, increases in IF output power closely correspond to increases in RF input power. Conversion compression occurs outside the linear range. The 1-dB compression point is where the conversion-gain is 1 dB lower than the conversion gain in the linear region of the mixer.

Mixer intermodulation-distortion (IMD) performance contributes to the limits on spur-free dynamic range. The highest possible values of second-order and third-order intercept points provide the greatest spur-free dynamic range.

The LO power coupled into the mixer controls performance. Inadequate LO power for a given mixer degrades conversion-gain and noise figure and, therefore, system sensitivity.

Conversion gain factor is specified at a particular LO drive level and is defined as the ratio of the numeric single-sideband (SSB) IF output-power to the numeric RF input-power. The ratio is converted to dB using equation (7).

(7) $$CG_{dB} = 10\log\left(\frac{P_{IF}}{P_{RF}}\right) = 10\log(P_{IF}) - 10\log(P_{RF})$$

where:

P_{IF} is the numeric single-sideband IF output-power in Watts
P_{RF} is the numeric RF input-power in Watts

Equation (7) yields a positive value for an IF output power greater than the RF input power indicating conversion gain. A negative result occurs for conversion loss.

Image-reject mixers are formed from a pair of balanced mixers. Input signals are offset by 90° through a hybrid splitter and translated to the IF outputs (with the same 90° offset). The IF outputs are passed through an IF hybrid to separate the desired and image sidebands. Because of tight phase requirements for the IF hybrid, such mixers are usually limited in bandwidth.

Local Oscillator

The local oscillator is a reference signal required for mixer injection to facilitate frequency translation in the receiver system. The LO is a large signal that drives the mixer diodes or transistors into a nonlinear region, thereby allowing the mixer to generate fundamental frequencies along with harmonics and mixing terms. The local oscillator signal mixes with the desired RF signal to produce an IF. The sum (LO+RF) and difference (LO–RF or RF–LO) terms are output to the IF mixer port. The oscillator frequency is tuned to select a desired frequency to be down-converted to an intermediate frequency. Proper planning for spurious signals ultimately determines the frequency offset (IF) between the oscillator and desired RF frequency. Once the oscillator is tuned, a signal that has the same spacing as the IF frequency away from the LO frequency is down-converted and passed through the IF filter.

In most wireless receivers, LO signals are generated by synthesized sources consisting of voltage-controlled oscillators (VCOs) stabilized by a phase-lock loop (PLL). A PLL consists of a phase detector, amplifier, loop filter, and VCO. In a PLL synthesizer, the VCO is locked in phase to a high-stability reference oscillator (usually a crystal oscillator). A phase detector compares the phase of a divided VCO frequency output to that of the precise reference oscillator and creates a correction voltage for the VCO based on phase differences between the reference and the VCO. A loop filter limits noise but also limits lock-time (frequency switching speed). Wider loops provide faster lock-times at the expense of higher reference spurs and phase noise. The loop filter must be narrow enough to limit oscillator reference spurs but wide enough to allow signal phase locking within the loop bandwidth with a tuning speed required by the receiver. The loop acquisition time is directly dependent on the time constant of the loop filter.

Key specifications for an LO (a receiver may have more than one, depending upon the number of IFs and the system architecture) include tuning range, frequency stability, spurious output levels, lock-time, and phase noise. Most of these specifications determine an LO's suitability for a particular wireless receiver application; the spurious and phase-noise performance also impact sensitivity and dynamic-range performance.

A noise analysis of a PLL oscillator must include several noise sources, including the VCO, the phase detector, and any prescalers or dividers used in the loop. In addition, noise sources on the VCO voltage control line, such as power-supply deviations and harmonics, can lead to spurious signals at the output of the PLL source.

IF Filter

An IF filter is a network that allows only an IF frequency to pass to the detection circuitry. System noise bandwidth is defined by the IF filter. If several IF filters are cascaded throughout the receiver line-up, the composite IF response determines system noise bandwidth. System noise bandwidth is key in determining a receiver's sensitivity level. In addition, IF filters reject signals very close to the desired signal. Adjacent channel selectivity is performed entirely by IF filtering in a superheterodyne receiver. (However, phase noise from the local oscillator can degrade adjacent channel selectivity.)

Ripple is the amount of amplitude variation induced by a filter on signals through the passband. Group delay describes the relative changes in phase of signals at different frequencies. One definition of group delay distortion is the difference between the maximum and minimum group delay value over the occupied signal bandwidth. High group-delay distortion results in phase distortion of pulsed or multi-frequency signals as well as information losses when processing digitally modulated signals.

Group delay is defined as the change in phase with respect to frequency, defined mathematically by equation (8).

(8) $$\mathrm{Gdelay} = -\frac{d\phi}{d\omega}$$

In cellular receiver front ends, SAW devices are usually used for IF filtering.

IF Amplifier

RF signals are translated down in frequency because amplifying a low frequency signal is easier (less expensive and a more efficient use of current) than amplifying a high frequency signal. Filtering close-in undesired adjacent channel signals with high-Q ceramic and crystal filters is also easier. Information recovery is easier at the lower IF frequency than the original RF frequency. The IF amplifier provides the necessary gain to boost the IF signal to a level required by the detector or to an additional down-conversion stage.

IF amplifier stages have less effect on the overall receiver noise figure, although dynamic-range characteristics are important to receiver performance. As with RF amplifier stages, IF amplifier linearity is important to the overall receiver dynamic range since intermodulation products can limit system performance.

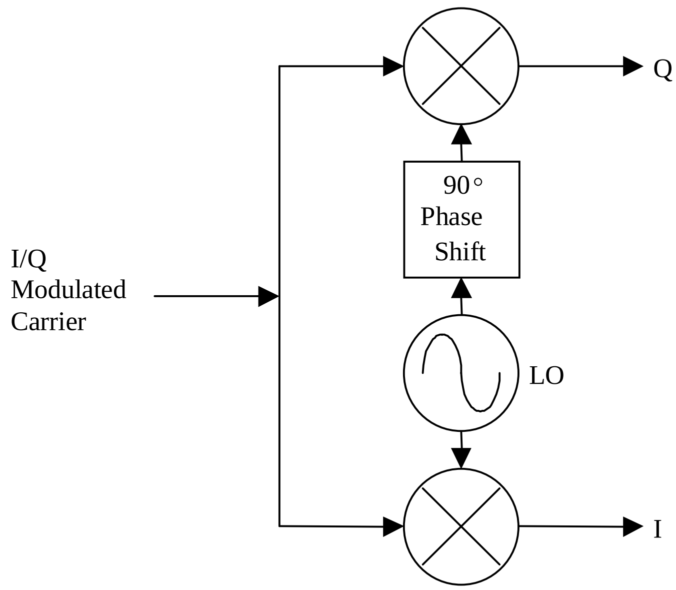
For optimum performance, IF amplifiers should be specified with the highest possible second- and third-order intercept points and adequate gain to drive the detector or additional down-conversion stages. However, interfering signals are usually greatly reduced by the IF filter, making the second- and third-order intercept point requirements of the IF amplifier minimal. Although noise figure is not as crucial in the IF amplifier stage, it can have an impact if gain in front of the stage is relatively low.

Detectors

A receiver detector is based on the type of modulation format. For simple FM, a quadrature detector offers a straightforward means of decoding modulation information. In this approach, deviations from resonant frequency cause deviations from phase quadrature, which is detected by a phase detector. To demodulate PSK signals, the amplitude of incoming signals must be limited and the phase detected.

An I/Q detector can detect and demodulate GMSK and other quadrature digital modulation formats. The I/Q demodulator is essentially a pair of double-balanced mixers offset by 90° and fed by a common, in-phase LO. Signals with π/4-DQPSK can be detected via coherent demodulation, differential detection, or frequency discrimination.

Figure 12. I/Q Detector



Receiver Sensitivity

Receiver systems are normally required to process very small signals. The weak signals can not be processed if the noise magnitude added by the receiver system is larger than that of the received signal. Increasing the desired signal's amplitude is one method of raising the signal above the noise of the receiver system. Signal amplitude can be increased by raising the transmitter's output power. Alternately, increasing the antenna aperture of the receiver, the transmitter, or both allows a stronger signal at the receiver input terminals. Increasing the physical size of an antenna is one method to increase its aperture.

Higher heat dissipation is typically required to increase transmitter output power. Cost, government regulations, and interference with other channels also limit the transmitter power available for a given application. Physically increasing transmitter antenna size may cause weight and wind load problems on the tower to which the antenna is mounted. Increasing receiver antenna size obviously increases the housing size for a portable product enclosing the antenna structure, such as a cellular telephone or pager.

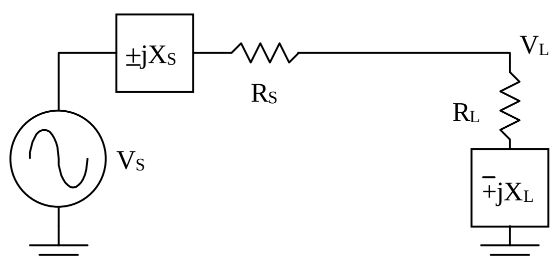
Because raising the desired signal amplitude above the noise added by the receiver may not always be practical, a weak signal might be processed by lowering the added noise. In this case, the noise must be decreased such that the noise amplitude is somewhat below the weak signal amplitude.

Noise of Two-Port Networks

To reduce noise added by a receiver system, the underlying causes of noise must be evaluated. Broadband noise is generated and subsequently categorized by several mechanisms, including thermal noise and shot noise. Other causes include recombination of hole/electron pairs (G-R noise), division of emitter current between the base and collector in transistors (partition noise), and noise associated with avalanche diodes. Noise analysis is based on available power concepts.

Available Power: The power a source would deliver to a conjugately matched load is defined as the *available power* from the source (maximum power transferred). Half the power is dissipated in the source and half the power is dissipated (transmitted) into the load under these conditions. The amount of power delivered is easily determined, as shown in Figure 13, which connects a complex source to a complex conjugate load termination.

Figure 13. Complex Source Connected to Complex Conjugate Load Termination



Equations (9) through (11) determine the available power.

(9) $$\text{IF} \quad Z_S = Z_L^* \quad \text{then} \quad Z_S = R_S \pm jX_S \quad \text{and} \quad Z_L = R_L \mp jX_L$$

For conjugate matched conditions, the voltage V_L is half of V_S. Using $R_L = R_S$, the power dissipated in R_L is easily calculated. Since

(10) $$V_L = \frac{V_S}{2}$$

then

(11) $$P_L = \frac{V_L^2}{R_L} = \frac{V_S^2}{4R_L} = \frac{V_S^2}{4R_S} = P_{Available}$$

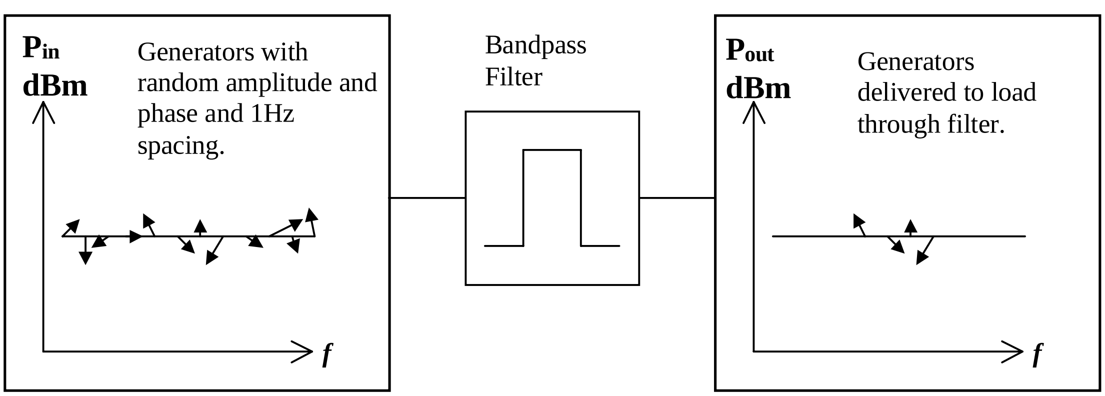
Thermal Noise: Simply stated, thermal noise is generated by vibrating conduction electrons. The amount of vibration, or kinetic energy, is proportional to the ambient temperature. These vibrations have broadband spectral content and occupy the desired signal bandwidth. The noise power available from a particular body, such as a resistor, would be that power that could be delivered to a load of ideal equivalent resistance. The ideal equivalent resistor would not produce thermal noise. Power is defined as the rate of energy removed or dissipated. Available power is the maximum rate at which energy is removed from a source and is expressed in joules/second. One joule/second is equivalent to one watt. Available thermal noise power from any loss is computed by taking the product of Boltzmann's constant, absolute ambient temperature, and bandwidth of the transmission path.

Boltzmann's constant, k, is the average mechanical energy per particle that can be coupled out electrically per degree of temperature. Boltzmann's constant is related to the universal gas constant R and is used in the ideal gas law, which states that PV = nRT. R gives the energy per mole of gas per degree, whereas k gives the average energy per particle per degree. The ratio R/k is equal to the number of particles in a mole, which is Avogadro's number of 6.02×10^{-26}. Boltzmann's constant k is 1.38×10^{-23}J/°K and is a conversion constant that converts energy in terms of absolute temperature or in terms of joules.

As ambient temperature increases, electron vibration also increases, thus causing an increase in available noise power. Absolute temperature is expressed in degrees Kelvin.

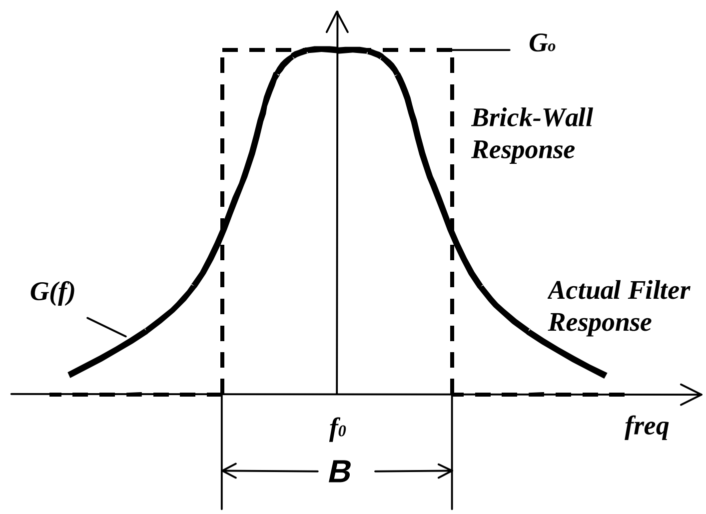
Thermal noise is independent of frequency for the most part. Thus, the available noise power at a specific discrete frequency is the same as the available noise power at some other specific discrete frequency. Thermal noise may be visualized as an infinite number of very small signal generators covering all frequencies, where each generator has a random amplitude and phase, as shown in Figure 14. If a bandpass filter with a 1-Hz bandwidth is centered at a specific discrete frequency and included in the transmission path, the noise power available from that discrete frequency is all that is delivered to an ideal matched load. The filter rejects all other frequencies. As the bandwidth of the filter is increased, more signal generators are coupled to the ideal load, resulting in an increase in total available noise power. Thus, the noise power available from any loss is directly proportional to the bandwidth of the transmission path.

Figure 14. Thermal Noise Visualized as Signal Generators



Equivalent noise bandwidth assumes a brick-wall filter response. A brick-wall response has 0 dB of insertion loss in the passband and infinite attenuation in the stopband. Receiver systems do not have ideal brick-wall filter responses, although some may approach the ideal case. The filter's 3-dB bandwidth is not the equivalent noise bandwidth. Noise power in the filter's skirts contributes to the noise power delivered to the load. The filter response is equated to the brick-wall filter response for noise power determination, as shown in Figure 15.

Figure 15. Filter Response Equated to Brick-Wall Filter Response



The area under the filter response curve is found by integrating the filter transfer function (response) as a function of frequency and then dividing by the maximum gain G_0 at frequency f_0 using equation (12). The integrand gives the equivalent noise bandwidth to the brick-wall response (see Figure 15).

(12) $$B = \int_0^{\infty} \frac{G(f)}{G_0} df$$

where:

B is the equivalent noise bandwidth.
G(f) is the transfer function versus frequency.
G_0 is the numeric gain at the reference frequency f_0, (usually at maximum gain).

If several filters are cascaded together, the equivalent noise bandwidth is obtained from the composite response.

Thermal noise power available from any loss, in equation form, is:

(13) $$P_W = kTB$$

where:

k ≡ Boltzmann's constant = 1.38×10^{-23} J/°K.
T ≡ absolute temperature.
B ≡ equivalent noise bandwidth.

This equation allows available noise power calculations for specific temperatures and noise bandwidths. For a reference temperature T_0 of 290°K and a 1-Hz equivalent noise bandwidth, available thermal noise power is expressed mathematically as

(14) $$P_W = kT_0B = 4.002 \times 10^{-21} \text{Watt}$$

Converting from watts to dBm yields

(15) $$P_{dBm} = -174 \text{ dBm}$$

where:

$T_0 \equiv$ absolute reference temperature = 290°K.
$B \equiv$ equivalent noise bandwidth = 1 Hz.

The reference temperature of 290°K and noise bandwidth of 1 Hz are utilized to calculate noise floor and sensitivity as discussed in later sections.

Shot Noise: The quantized and random nature of current flow generates shot noise. Current flow is not continuous but is limited by an electron's unit charge ($e = 1.6 \times 10^{-19}$ coulombs). Current flow across a given boundary is quantized with a particular number of electrons, or holes, crossing the boundary at a given instant. Charged particles also flow with random spacing. At any given instant, the number of charged particles flowing across a boundary vary around some average value. When a DC current flows, the average current is the measured DC value. The measurement does not yield current variation information or associated frequency information. Statistical analysis of the random occurrence of particle flow shows the mean square current variations are uniformly distributed in frequency and have a spectral density mathematically defined by:

(16) $$\overline{i_n^{\,2}(f)} = 2eI_0B \, {}^{A^2}\!/\!_{Hz}$$

where:

I_0 is the DC current.
$e \equiv$ electron charge = 1.6×10^{-19} coulomb.

Available Gain: Available gain G_a for a two-port is defined as the ratio of power available at the output to the power available from the source, mathematically expressed by equation (17).

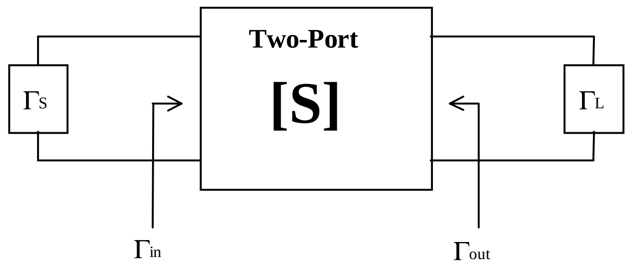
(17) $$G_a = \frac{P_{ao}}{P_{as}}$$

where:

Pao is the available output power.
Pas is the power available from the source.
Both powers in the ratio are expressed numerically.

Figure 16 shows a two-port network connection for available gain definition. Γ_{in} is the two-port input reflection coefficient and the output is terminated in Γ_L. Γ_{out} is the two-port output reflection coefficient as long as the input is terminated with Γ_S.

Figure 16. Two-Port Network Connection for Available Gain Definition



Available gain factor depends on the source termination Γ_S and the s-parameters of the two-port, as shown by equation (18). Available gain dependence on Γ_S may cause two-port mismatch at the input. Available gain depends on the two-port output being conjugately matched to Γ_L; thus, $\Gamma_L=\Gamma_{out}^*$. Γ_{out} is a function of Γ_S and the two-port s-parameters shown by equation (20).

(18) $$G_a = \frac{|S_{21}|^2\left(1-|\Gamma_S|^2\right)}{\left(1-\left|\frac{S_{22}-(\Delta)\Gamma_S}{1-S_{11}\Gamma_S}\right|^2\right)|1-S_{11}\Gamma_S|^2}$$

(19) $$\Delta = S_{11}S_{22} - S_{21}S_{12}$$

(20) $$\Gamma_{out} = S_{22} + \frac{S_{21}S_{12}\Gamma_S}{1-S_{11}\Gamma_S}$$

The power available to the two-port output load is a function of power delivered to the two-port input. The power delivered to the two-port input depends on mismatch between Γ_S and Γ_{in}. Determining the power available from the source requires a complex conjugate load termination on the source. The noise power available from the source is kT_SB from equation (13), where T_S is the source noise temperature. Equation (18) computes available gain G_a numerically, utilized in subsequent derivations.

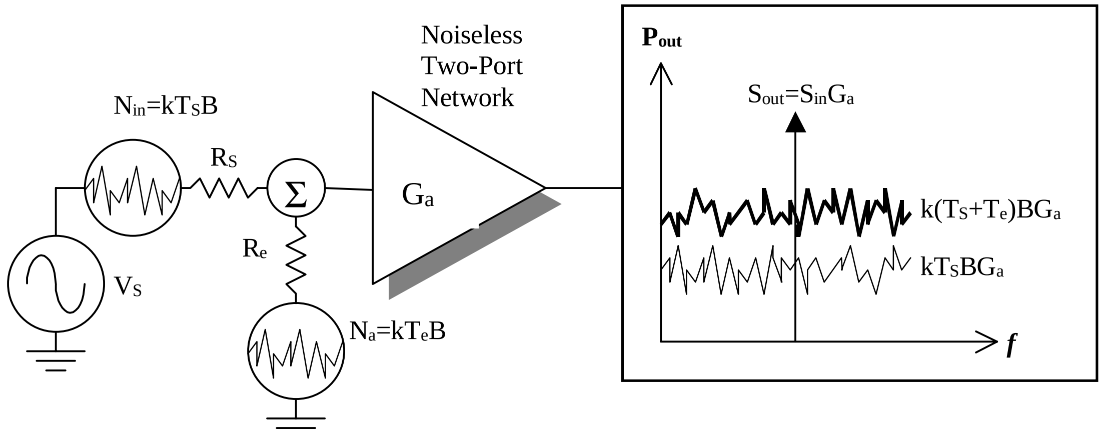
Noise Temperature

Thermal noise power is generated in source losses connected to the two-port input. This noise is described in terms of a noise temperature T_S. (In this derivation, it is assumed that the thermal noise from the source does not contain phase noise.) Available source noise power N_{in} is defined in terms of noise temperature T_S and noise bandwidth by equation (21).

(21) $$N_{in} = kT_SB$$

N_{in} is the noise power available to the input.
T_S is the source noise temperature in °Kelvin.

Figure 17. Available Thermal Noise (Noise Temperature)



The available thermal noise N_{in} is represented by a noise voltage source in series with the source resistor R_S, as shown in Figure 17. If the two-port network is noiseless, the noise delivered to the output N_{out} is the source noise kT_SB times the available gain G_a. The input signal S_{in} is represented by the AC voltage source in series with the source resistor R_S as shown. The signal available to the load is the input signal S_{in} times the available gain G_a. Equations (22) and (23) calculate noise N_{out} and signal S_{out} delivered to the output respectively for the noiseless two-port.

(22) $$N_{out} = kT_SBG_a$$

(23) $$S_{out} = S_{in}G_a$$

Noise in electronic two-port circuits is generated from many sources, such as thermal noise, shot noise, and partition noise. Noise characterization of two-port networks typically combines all internally generated noise sources and assumes all are due to thermal noise. The internally generated noise sources are removed from a noisy two-port, thus creating a noiseless two-port with an additional thermal noise generator kT_eB connected to the input (see Figure 17).

An input effective noise temperature T_e is calculated based on the two-port's added noise power. This description does not mean the device is at a particular noise temperature, but that the noise power available is equivalent to a thermally passive resistive source at that temperature. Externally generated interference is not included in the noise temperature definition. Lightning, sparks, arcing, or spurious responses are examples of external interference. Noise temperature T_e is mathematically defined in terms of available noise power of internal two-port noise sources and noise bandwidth in equation (24).

(24) $$T_e = \frac{N_a}{kB}$$

where:

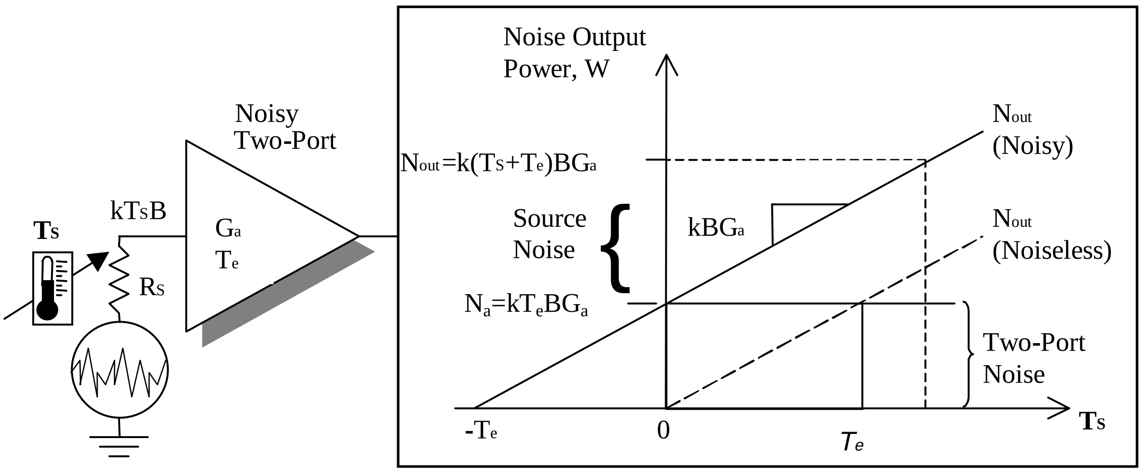
T_e is the two-port equivalent input noise temperature in °K.
N_a is the added noise power due to two-port internal sources.

Adding source noise N_{in} and added noise N_a and multiplying the result by the available gain G_a determines the noise power N_{out} delivered to the matched load for the noisy two-port. Equation (25) determines the noise delivered to the output.

(25) $$N_{out} = k(T_S + T_e)BG_a$$

Figure 18 is a plot of noise output power versus source temperature. If a resistor is connected to the input of a noiseless two-port network, the noise available to the output load is given by equation (22).

Figure 18. Noise Output Power Versus Source Temperature (Noise Temperature)



If the ambient source temperature varies from 0°K in an upward direction, the noise power available to the load varies in a straight line according to equation (22), as shown in Figure 18. (This result assumes the two-port is not compressed.) The noiseless two-port output noise power characteristic intersects the vertical axis (y-intercept) at the origin as shown. This indicates zero noise power with a source temperature of 0°K since the two-port adds no noise of its own. The noise characteristic's slope is a function of the two-port's gain-bandwidth product, more specifically kBG_a.

For the noiseless two-port, noise power available to the load is simply the noise available from the source multiplied by the available gain factor of the two-port, mathematically kT_SBG_a (equation (22)). Noise delivered to a matched load for the noiseless two-port is found by graphically extending a vertical line from the source temperature T_S. A horizontal line is then drawn where the vertical T_S line intersects the noiseless gain-bandwidth product diagonal. Noise power available at the output is determined from the horizontal line intersection with the vertical axis.

If a source resistor is connected to the input of a noisy two-port network, the noise delivered to a matched load is given by equation (25). Allowing the ambient source temperature to vary from 0°K in an upward direction produces a straight-line noise characteristic as previously described but with a higher y-intercept, (see Figure 18). The noiseless two-port linear noise characteristic shifts upward by the added noise of a noisy two-port, which gives a y-intercept of kT_eBG_a.

After extrapolating the noise characteristic downward for the noisy two-port, the diagonal passes through the x-axis at the negative of T_e. The slope is the same as that of the noiseless two-port, specifically kBG_a. Extending a horizontal line through the y-intercept of the noisy characteristic intersects the noiseless characteristic, as depicted by a solid horizontal line in Figure 18. The point intersects at a source temperature equal to the noise temperature T_e of the noisy two-port.

The noisy two-port noise characteristic determines output noise based on a source temperature. If the source temperature is set to 0°K, the only noise power delivered to the load would be the noise generated internally by the two-port, strictly kT_eBG_a.

In summary, the noise power delivered to a noisy two-port's matched load is the sum of the source noise and the added noise referenced to the input, multiplied by the two-port's available gain factor, mathematically $(kT_SB+kT_eB)G_a$ (see equation (25)). Noise delivered to the noisy two-port's matched load is graphed by extending a vertical line from the source temperature T_S (a dashed vertical line in Figure 18). A horizontal line is drawn where the vertical T_S line intersects the noisy gain-bandwidth product diagonal (a dashed horizontal line in Figure 18). Noise power available at the output is represented by the horizontal line intersection with the vertical axis.

Noise temperature is one method for specifying the noise generated in electronic networks. This method is most commonly used in space or satellite applications where source temperatures deviate from 290°K. Noise temperature characterization does not require a specific source temperature, as is the case for noise figure.

Noise Figure

Noise factor definition sets the source temperature to a reference of 290°K, thereby providing a specific amount of noise power available from the source, varied only by the noise bandwidth. Source noise power N_{in} is defined in terms of reference noise temperature T_0 and noise equivalent bandwidth depicted in equation (26).

(26) $$N_{in} = kT_0B$$

Figure 19. Available Thermal Noise (Noise Figure)

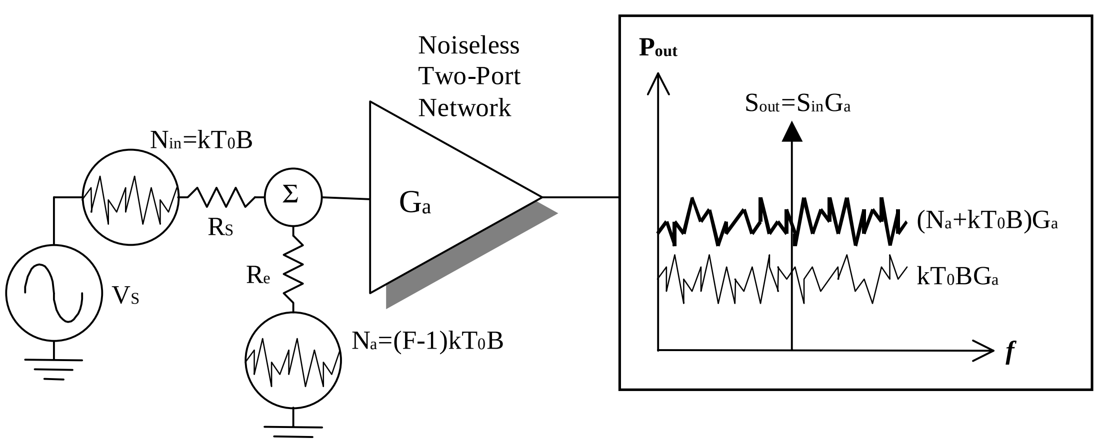


Figure 19 shows the available thermal noise N_{in} represented by a noise voltage source in series with the source resistor R_S. Noise factor is defined when the source temperature T_S is set to $T_0 = 290°K$. If the two-port is noiseless, the noise delivered to the output N_{out} is the source noise kT_0B times the available gain G_a. The input signal S_{in} is represented by the AC voltage source in series with the source resistor R_S as shown. The signal delivered to a conjugately matched load S_{out} is the input signal S_{in} times the available gain G_a. Equations (27) and (28) compute noise N_{out} and signal S_{out}, respectively, delivered to the output, when the source temperature is T_0.

(27) $$N_{out} = kT_0BG_a$$

(28) $$S_{out} = S_{in}G_a$$

Removing the internally generated noise sources from the noisy two-port creates a noiseless two-port with an additional thermal noise generator N_a connected to the input as shown in Figure 19. The added noise is based on the two-port noise factor and is summed with the source noise at the input. Equation (29) mathematically defines the two-port's noise power added, N_a , in terms of noise factor and source noise power at the reference temperature T_0.

(29) $$N_a = (F-1)kT_0B$$

where:

N_a is the added noise power due to two-port internal sources.
F is the two-port noise factor.

Noise from the source N_{in} and noise from the added noise generator N_a are summed and multiplied by the available gain G_a to determine the noise power delivered to the load N_{out}. Equation (30) determines the noise delivered to the noisy two-port's output.

(30) $$N_{out} = (N_a + kT_0B)G_a = FkT_0BG_a$$

Figure 20. Noise Output Power Versus Source Temperature (Noise Figure)

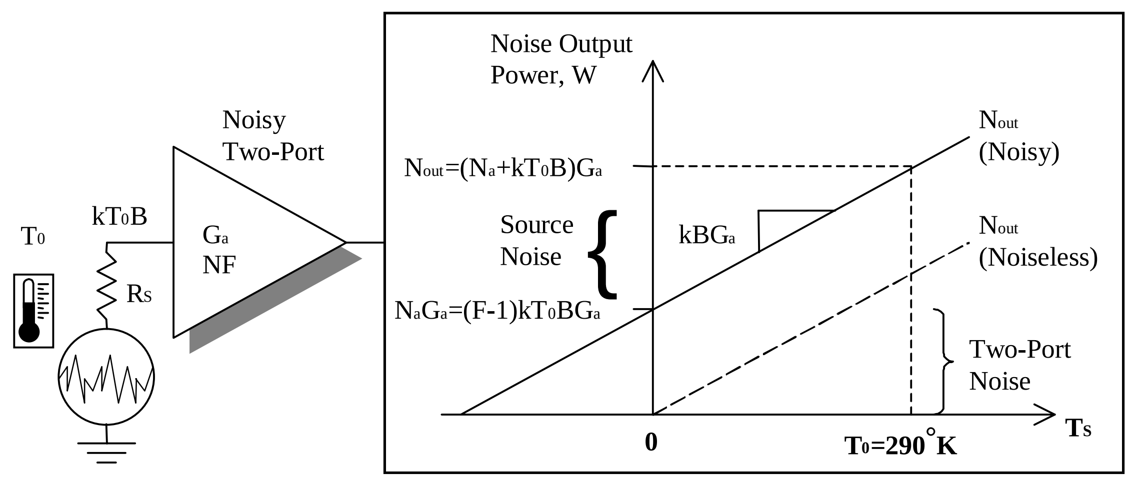


Figure 20 plots noise output power versus source temperature. If a resistor is connected to the input of a noiseless two-port network, the noise available to the output load is given by equation (22). Varying ambient source temperature from 0°K upward produces noise power with a straight-line characteristic according to equation (22), as shown in Figure 20. The noiseless two-port output noise power characteristic intersects the vertical axis (y-intercept) at the origin as shown. Since the two-port adds no noise of its own, this result indicates zero noise power with a source temperature of 0°K. The slope of this noise characteristic is kBG_a.

If the two-port adds noise, the output noise varies in a straight line, as previously described, with a linear sweep of source temperature. The y-intercept is given in terms of the noise factor and the noise bandwidth, as illustrated in Figure 20 and defined by equation (31).

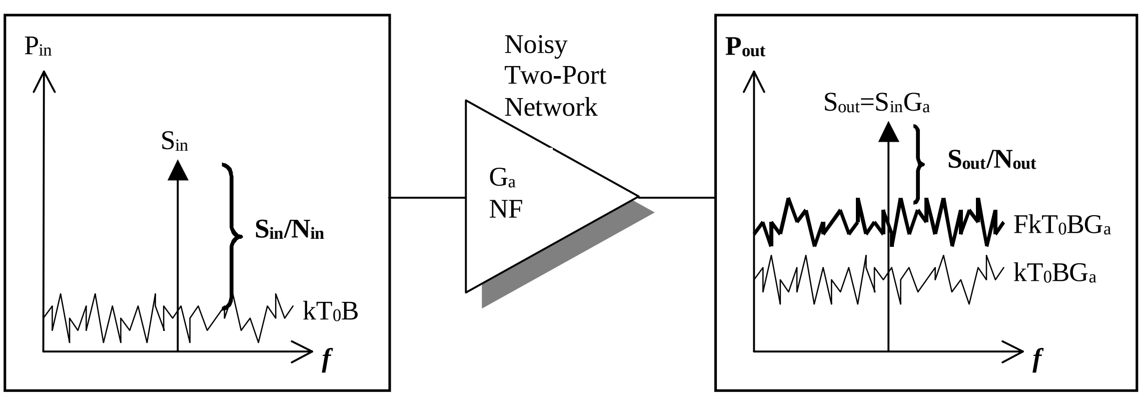
(31) $$N_a G_a = (F-1)kT_0 BG_a$$

With a source temperature set to 0°K the only noise power delivered to the load is the noise generated internally by the two-port, strictly $(F–1)kT_0BG_a$. The kBG_a slope is the same as that slope for the noiseless two-port temperature sweep. The noisy response with temperature is therefore parallel to the noiseless response. The noisy two-port noise characteristic determines output noise based on a source temperature of 290°K as described by equation (30).

In summary, the noise power N_{out} delivered to a matched load for a noisy two-port is the product of two-port noise factor, source noise, and available gain factor, mathematically FkT_0BG_a (see equation (30)). Noise delivered to a matched load for the noisy two-port is graphed by extending a vertical line from the source temperature T_0 (a dashed vertical line in Figure 20). A horizontal line is constructed where the vertical T_0 line intersects the noisy gain-bandwidth product diagonal (a dashed horizontal line in Figure 20). The horizontal line intersection with the vertical axis represents output noise power N_{out}.

Noise figure (noise factor) quantifies signal-to-noise ratio degradation as a signal is processed through a network, as shown in Figure 21. The network has at least two ports—an input and an output. Because the terms *noise figure* and *noise factor* are sometimes used interchangeably, omitting the dB designation may cause confusion. Noise factor is a numerical value and typically represented with a capital letter F. Noise figure is the dB representation of the noise factor and denoted with capital letters NF. Signal and noise powers in Figure 21 are portrayed numerically.

Figure 21. Noise Figure Quantifies Signal-to-Noise Ratio Degradation



Noise factor is defined mathematically as the numeric signal-to-noise ratio at the input divided by the numeric signal-to-noise ratio at the output. The definition is determined at a source reference temperature T_0 of 290°K ($T_S = T_0 = 290$°K). In equation form:

(32)
$$F = \left.\frac{S_{in}/N_{in}}{S_{out}/N_{out}}\right|_{T_S=T_0=290°K} = \left.\frac{N_{out}}{G_a N_{in}}\right|_{T_S=T_0=290°K}$$

Noise figure is noise factor converted to dB (at $T_S = T_0 = 290$°K). In equation form:

(33)
$$NF = 10\log(F) = 10\log\left(\frac{S_{in}/N_{in}}{S_{out}/N_{out}}\right) = 10\log\left(S_{in}/N_{in}\right) - 10\log\left(S_{out}/N_{out}\right)$$

where:

S_{in} is numeric input signal power.
N_{in} is numeric input noise power.
S_{out} is numeric output signal power.
N_{out} is numeric output noise power.

The signal at the output S_{out} is simply the input signal S_{in} times the numeric available gain, G_a, thus $S_{out} = S_{in}G_a$. If the amplifier did not add noise, the output noise power N_{out} would be the noise power N_{in} times the numeric available gain G_a. Because the amplifier does add noise, additional noise power N_a exists at the output. The total noise at the output is therefore $G_a(N_{in}+N_a)$. Equation (32) becomes equation (34).

(34)
$$F = \frac{S_{in}/N_{in}}{S_{out}/N_{out}} = \frac{S_{in}/N_{in}}{S_{in}G_a/G_a(N_{in}+N_a)} = \frac{N_{in}+N_a}{N_{in}}$$

Equation (34) shows that the noise figure is a function of the added noise N_a and source noise N_{in}. Thermal noise power at the input changes with ambient temperature. Thermal noise power is equated with a source temperature. The IRE (now the IEEE) therefore defines a reference for noise factor (noise figure) at a source temperature T_0 of 290°K. Thermal noise power at the two-port's input at reference temperature is therefore kT_0B. Substituting kT_0B into equation (34) for N_{in} yields the IEEE definition of noise factor at a reference temperature of 290°K.

(35)
$$F = \frac{G_a(kT_0B + N_a)}{G_a kT_0B} = \frac{N_{in}+N_a}{N_{in}} = \frac{k(T_0+T_e)B}{kT_0B} = \frac{T_0+T_e}{T_0} = \frac{T_e}{T_0} + 1$$

where:

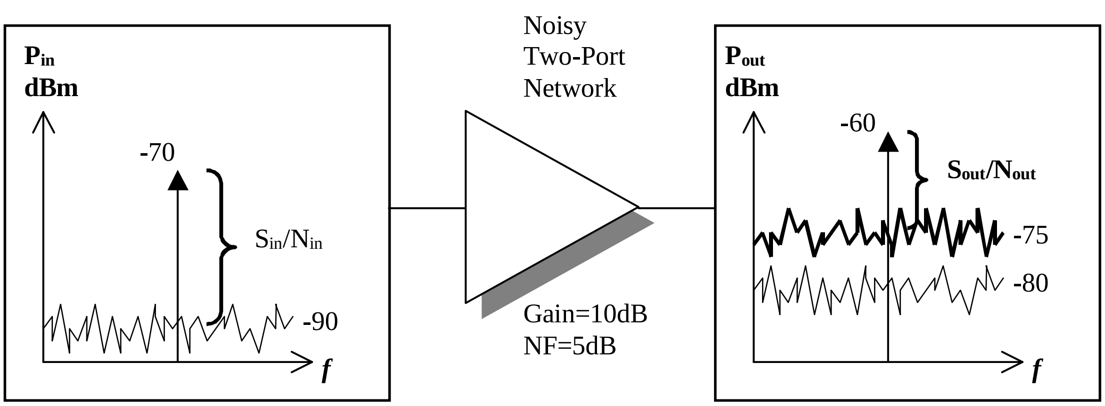
$T_0 \equiv$ reference temperature = 290°K.
$T_e \equiv$ input referenced two-port equivalent noise temperature.
$B \equiv$ equivalent noise bandwidth.
$G_a \equiv$ available gain of the two-port.
$N_a \equiv$ noise added by two-port referenced to the input.
$N_{in} \equiv$ source noise.

As equation (35) demonstrates, noise factor is independent of signal level. Furthermore, added noise power N_a is proportional to bandwidth as is the input noise (that is, $(F–1)kT_0B$). The bandwidth dependence in the numerator cancels the bandwidth dependence in the denominator of equation (35). Noise factor is, therefore, independent of bandwidth. Based on the IEEE Standard definition of noise factor, noise factor is the ratio of total noise power output to the portion of noise power output due to noise power at the input. The ratio is defined when the input source temperature is 290°K, as depicted in equation (35).

Figure 22 shows an amplifier with 10 dB of gain and a 5-dB noise figure. All signals and noise powers are expressed logarithmically. A –70-dBm signal is connected to the input of the amplifier. Because the amplifier has 10 dB of gain, a –60-dBm signal power is delivered to the load. The noise floor at the amplifier's input is –90 dBm due to thermal source noise. This example has a 20-dB signal-to-noise ratio at the input because the desired signal is 20 dB above the noise [–70 dBm – (–90 dBm)]. Both the noise at the amplifier's input and the desired signal are amplified. If the amplifier did not add its own noise, the noise power delivered to the output load would be –80 dBm. This yields a 20-dB signal-to-noise ratio at the output [–60 dBm – (–80 dBm)] with no signal-to-noise ratio degradation as the signal is processed.   The noiseless amplifier thus provides a noise figure of 0 dB.

However, the amplifier does add noise in addition to the amplified source noise, thereby delivering more noise to the load. The noise delivered to the matched output load is –75 dBm [–90 dBm + 10 dB + 5dB]. The noisy amplifier signal-to-noise ratio at the output is 15 dB [–60 dBm – (–75 dBm)] instead of the 20 dB reflected in the 0-dB noise figure case. Signal-to-noise ratio degrades because the amplifier adds noise. In this case, a degradation of 5 dB is noticed, hence, a 5-dB noise figure.

Figure 22. Amplifier With 10 dB of Gain and a 5-dB Noise Figure



Noise temperature is converted to noise factor using equation (36).

(36) $$F = \frac{T_e}{T_0} + 1$$

Noise temperature is derived from noise factor by rearranging equation (36) to get equation (37).

(37) $$T_e = (F - 1)T_0$$

Noise Figure Versus Noise Temperature: Noise figure is more commonly used for designing and analyzing earth based equipment partially because it is more intuitive than noise temperature. Noise figure values typically vary over a convenient range; for example, earth based equipment may vary between 1 dB and 15 dB. Noise figure is not appropriate for designing deep space applications since the source temperature deviates dramatically from 290°K, usually around ~4°K. More measurement error is encountered as the source temperature deviates from the specified 290°K, particularly for lower noise circuits. Using noise figure in system design and analysis requires more computation than using noise temperatures.

Noise temperature is most commonly used for designing and analyzing deep space applications. Noise temperature does not require a source reference temperature. Measurement uncertainties depend on the two-port noise temperature T_e and source temperature T_S. Noise temperature values are typically large numbers ranging from 75°K to 9000°K (for noise figures between 1 dB and 15 dB).

The objective of both noise figure and noise temperature is to determine how much signal-to-noise degradation an RF signal experiences as it processes through a network. The two-port network may consist of one or more two-ports connected (cascaded) together to form a complete system. If several stages with known gains and noise figures are cascaded together, the total system gain and noise figure determines the overall noise degradation. The system's total noise figure or noise temperature is used along with its noise bandwidth to determine the smallest signal the system processes. Total gain is calculated by adding the gain in dB of each individual stage, less any mismatch loss. Total noise figure or noise temperature requires more complex calculations.

Cascaded Noise Temperature

Receiver systems are typically configured by cascading several individual blocks together, with each block performing a specific function to process the desired signal. Most receivers perform three basic processing functions.

1) First, a desired signal, usually very weak, is amplified enough to recover information. The signal passes through several filters unimpeded, and amplification is distributed throughout the system.

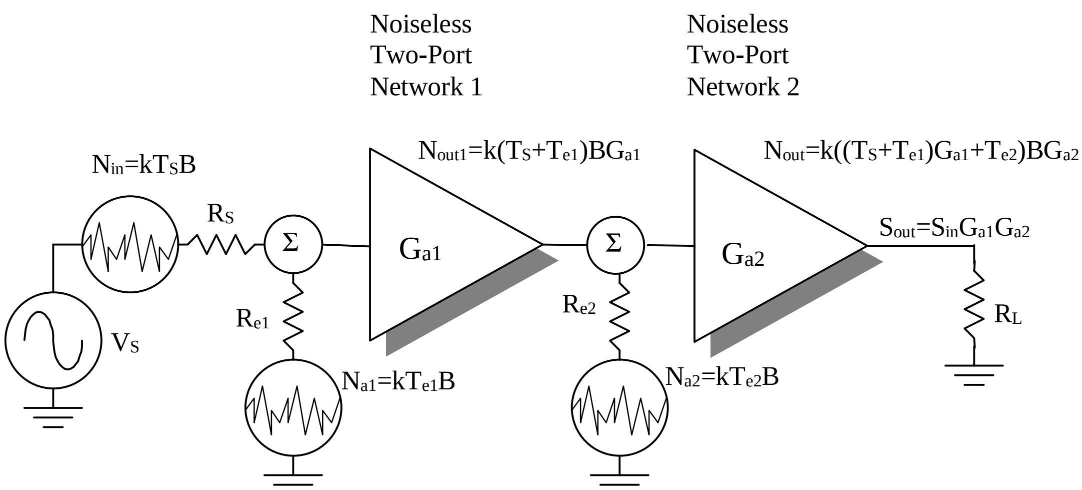
2) Next, all undesired interfering signals are rejected or filtered out. Several filters, distributed along the receiver chain attenuate a particular interfering signal.

3) Finally, the receiver demodulates (recovers) intelligible information from the RF modulated signal. While processing, the signal degrades due to noise.

Frequency translation known as mixing is also common but is not treated as a separate design or analysis function with respect to gain and noise figure. The mixing function is converted to a two-port response in a cascaded noise figure and noise temperature analysis.

A two-port network's available gain and noise temperature (noise figure) characterize the signal-to-noise ratio degradation of a processed RF signal. Cascaded blocks combine to form a new two-port, also described by gain and noise temperature (noise figure). Signal level is increased by gain stages and decreased by losses. A signal entering the receiver input is small and therefore comparable to noise levels at the input. Each system gain block increases the signal's amplitude such that it becomes larger with respect to noise magnitudes added by subsequent stages. Accordingly, noise added by the later system stages contributes less to the overall system noise figure. Consider the two-stage network shown in Figure 23.

Figure 23. Two-Stage Network (Cascaded Noise Temperature)



Two gain stages are cascaded together, where the added noise N_a of each stage is removed and referenced to the input. The output noise of the first stage is a combination of source noise kT_SB and added noise $kT_{e1}B$ multiplied by the available gain. The first stage output noise N_{out1} is given by equation (25), as shown in Figure 23. This noise combines with the added noise referenced to the input of the second gain stage $kT_{e2}B$. The second stage gain amplifies this combined noise, thus producing the total output noise shown in the figure and given by equation (38).

(38) $$N_{out} = \left(k\left(T_S + T_{e1}\right)BG_{a1} + kT_{e2}B\right)G_{a2} = k\left(\left(T_S + T_{e1}\right)G_{a1} + T_{e2}\right)BG_{a2}$$

N_{out} in equation (38) represents the sum of source noise N_{in} and combined two-port added noise Na times the available gain. Mathematically

(39) $$N_{out} = \left(N_{in} + N_a\right)G_{aT} = kT_SBG_{a1}G_{a2} + kT_{e1}BG_{a1}G_{a2} + kT_{e2}BG_{a2}$$

(40) $$\text{where: } G_{aT} = G_{a1} \times G_{a2}$$

Dividing equation (39) by the total available gain of equation (40) references noise to the input. Equation (39) becomes equation (41)

(41) $$\left(N_{in} + N_a\right) = kT_SB + kT_{e1}B + \frac{kT_{e2}B}{G_{a1}}$$

in which the input noise N_{in} is defined by equation (21) (kT_SB). The total noise added by the cascaded gain stages is determined by subtracting the source noise N_{in} from both sides of the equation.

(42) $$N_a = kT_SB + kT_{e1}B + \frac{kT_{e2}B}{G_{a1}} - kT_SB = kT_{e1}B + \frac{kT_{e2}B}{G_{a1}}$$

The cascaded two-port's total noise temperature is determined by converting added noise N_a to an equivalent noise temperature using equation (21). Thus, dividing equation (42) by kB yields the total equivalent noise temperature T_{eT} of the cascaded network given by equation (43).

(43) $$T_{eT} = T_{e1} + \frac{T_{e2}}{G_{a1}}$$

The cascaded noise temperature equation is extended to multi-stage networks.

Figure 24. Cascaded Receiver System (Cascaded Noise Temperature)

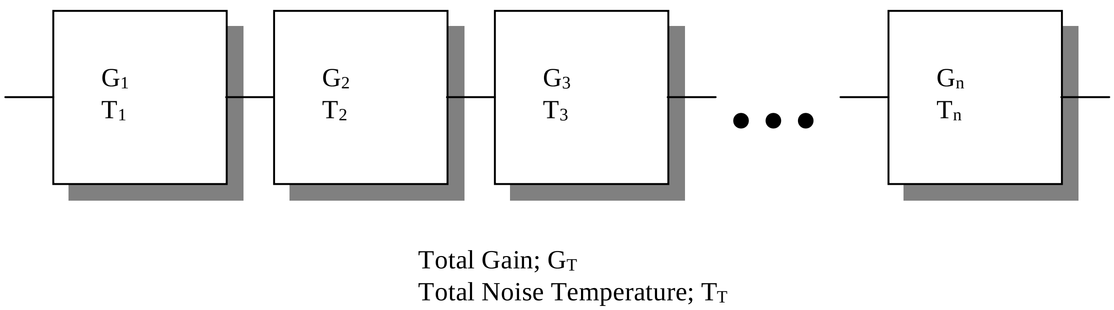


Figure 24 shows a system with several cascaded stages forming a receiver system. This system has an *n*-number of stages in which both available gain and noise temperature are known for each stage. Some of the blocks may have gain, whereas others have loss. Available gains expressed numerically are greater than unity for stages that increase the signal level and less than unity for stages that decrease the signal level. A signal level unchanged in amplitude is equivalent to unity gain. Each stage is sequentially numbered with subscripts, referenced in equations (44), (45), and (46) and in Figure 24. All gains are defined as the available gain of each stage. The total noise temperature T_T of the cascaded system is given by equation (44).

(44) $$T_T = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \cdots + \frac{T_n}{G_1G_2 \cdots G_{n-1}}$$

where:

T_1, T_2, T_3,…T_n are equivalent noise temperatures.
G_1, G_2, G_3,…G_{n-1} are available gain factors.

Total numeric available gain G_{aT} is defined mathematically by equation (45).

(45) $$G_{aT} = G_1 \times G_2 \times G_3 \times ... \times G_n$$

where:

G_1 is the first stage numeric available gain,
G_2 is the second stage numeric available gain, …
G_n is the n^{th} stage numeric available gain.

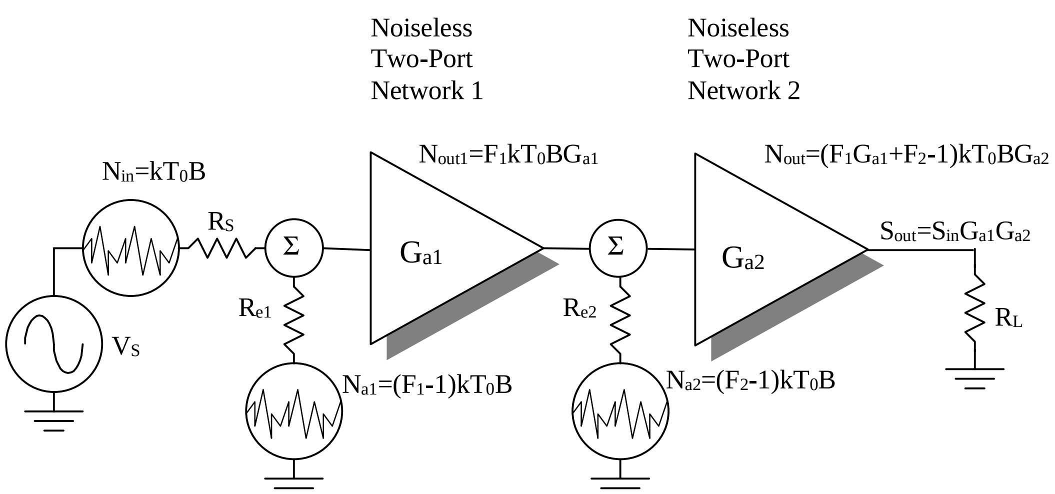
Total gain is converted to dB using equation (46).

(46) $$G_{dB} = 10\log(G_{aT})$$

Cascaded Noise Figure

Total system noise figure is similar to total noise temperature. Two gain stages are cascaded together where the added noise N_a of each stage is removed and referenced to the input, as shown in Figure 25. The output noise of the first stage is a combination of source noise at 290°K (kT_0B) and added noise $(F_1–1)kT_0B$ multiplied by the available gain G_{a1}. The first stage output noise N_{out1} is given by equation (30), as shown in Figure 25. This noise combines with added noise referenced to the input of the second gain stage $(F_2–1)kT_0B$. This combined noise is amplified by the second stage gain G_{a2}, thus producing the total output noise shown in Figure 25 and given by equation (47).

Figure 25. Two-Stage Network (Cascaded Noise Figure)



(47) $$N_{out} = F_1kT_0BG_{a1}G_{a2} + (F_2 - 1)kT_0BG_{a2} = (F_1G_{a1} + F_2 - 1)kT_0BG_{a2}$$

N_{out} in equation (47) represents the product of source noise at 290°K, total noise factor, and total available gain. (Referring to equation (30), N_{out} for a noisy two-port is the product of noise factor, source noise at 290°K and available gain.) Mathematically, the combined two-port output noise is determined by equation (48).

(48) $$N_{out} = F_TkT_0BG_{aT} = \left(F_1 + \frac{F_2 - 1}{G_{a1}}\right)kT_0BG_{a1}G_{a2}$$

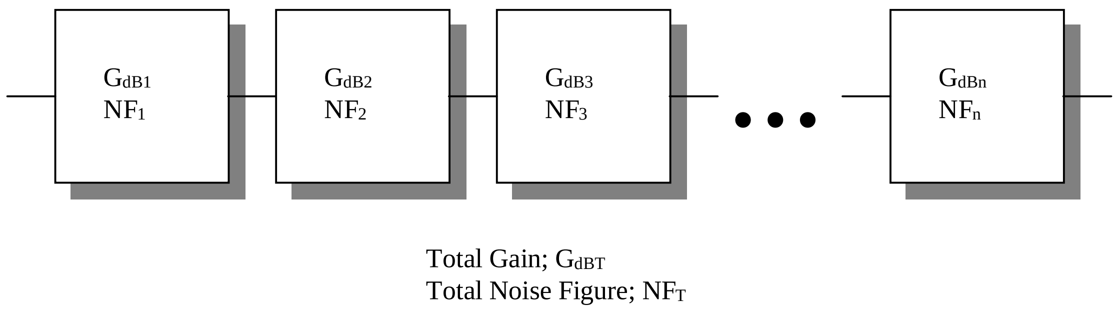
(49) $$\text{where: } G_{aT} = G_{a1} \times G_{a2}$$

Dividing equation (48) by the total available gain of equation (49) and the source noise kT_0B yields total noise factor. Equation (48) becomes equation (50).

(50) $$F_T = F_1 + \frac{F_2 - 1}{G_{a1}}$$

The cascaded noise factor equation applies to multi-stage networks. Figure 26 shows a system with several cascaded stages constituting a receiver system. This system has an *n*-number of stages in which both available gain and noise figure are known for each stage. Gain is expressed in dB for this system. Some of the blocks may have gain, whereas others have loss (negative gain). If a stage contributes a loss to the signal, the gain in dB is negative. A stage that increases the signal level is treated as a positive gain in dB. Each stage is sequentially numbered with subscripts referenced in the following equations. All gains are defined as the available gain of each stage. Harold Friis defined total system noise factor F_T mathematically by equation (51) conforming to Figure 26.

Figure 26. Cascaded Receiver System (Cascaded Noise Figure)



(51) $$F_T = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_n - 1}{G_1 G_2 \cdots G_{n-1}}$$

where

(52) $$F_1 = 10^{NF_1/10} \quad \cdots \quad F_n = 10^{NF_n/10}$$

(53) $$G_1 = 10^{G_{dB1}/10} \quad \cdots \quad G_{n-1} = 10^{G_{dB(n-1)}/10}$$

therefore: F_1, F_2, F_3,…F_n are noise factors.
G_1, G_2, G_3,…G_{n-1} are available gain factors.

The gain of the very last stage is not utilized in equation (51) and is therefore not converted to a numeric gain factor in equation (53). Equation (51) indicates noise factors of the preliminary stages contribute more to system noise factor than following stages. Further, additional gain in front of noisy two-ports lessens noise contributions of those circuits. However, too much gain in the front drives following stages into compression or distortion.

System blocks at the back of the chain contribute much less noise than blocks in the front. To optimize system noise factor, the noise factor of each stage is reduced as much as possible. A trade-off between noise and distortion (and compression) is made when adding additional gain in the front to take over noise in the back. Equation (54) demonstrates the conversion of system noise factor to total noise figure.

(54) $$NF_T = 10\log(F_T)$$

Total system gain G_{dbT} is defined as the sum of each individual block available gain G_a in dB expressed mathematically by equation (55.

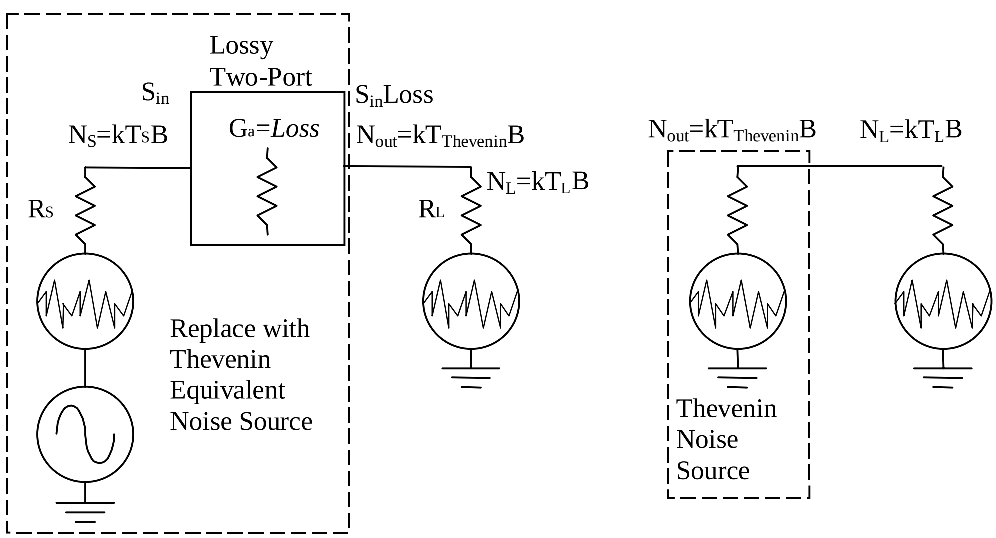
(55) $$G_{dBT} = G_{dB1} + G_{dB2} + G_{dB3} + ... + G_{dBn}$$

G_{dB1} is the first stage available gain in dB.
G_{dB2} is the second stage available gain in dB, …
G_{dBn} is the n^{th} stage available gain in dB.

Lossy Two-Port Noise Figure

Figure 27 shows a two-port that adds loss to a receiver system, such as a filter, attenuator, or circulator. The numeric power delivered to the load is the product of S_{in} and the numeric loss of the two-port (S_{in}Loss). The available gain factor G_a has a value less than unity, given by equation (56).

Figure 27. Lossy Two-Port (Noise Figure)



(56) $$G_a = \frac{S_{out}}{S_{in}} = \frac{S_{in}\,\text{Loss}}{S_{in}} = \text{Loss}$$

The noise delivered to a matched load is the thermal noise generated by the Thevenin equivalent of the original source and the lossy two-port. Mathematically,

(57) $$N_{out} = kT_{Thevenin}B$$

The noise temperatures for the source, load, and two-port network are equivalent. The load also generates thermal noise power that is available to the lossy two-port network at its output (kT_LB). The source, load, and two-port network are in thermal equilibrium with each other. Therefore, thermal noise available to the load from the two-port network and the source is equivalent to the noise available to the load from the source by itself. The noise power available to the two-port network's output from the load termination is equivalent to the noise power available from the source.

(58) $$kT_{Thevenin}B = kT_SB = kT_LB$$

In summary, the noise power available from the source is equivalent to the noise power available from the load, and the noise power available from the source connected to the lossy two-port is equivalent to the noise power available from the load.

Noise factor is defined as the numeric signal-to-noise ratio at the input of a network divided by the numeric signal-to-noise ratio at the output shown by equation (59).

(59) $$F = \frac{S_{in}/N_{in}}{S_{out}/N_{out}}$$

The available gain is given by equation (56), thus the reciprocal of gain is given by equation (60).

(60) $$\frac{1}{G_a} = \frac{S_{in}}{S_{out}}$$

Substitution of equation (60) into equation (59) gives noise factor in terms of gain and numeric noise powers.

(61) $$F = \frac{N_{out}}{G_aN_{in}}$$

Equation (56) mathematically defines the available gain to be equal to numeric loss for a lossy two-port network $(G_a = \mathrm{Loss})$. Substituting numeric loss for the available gain into equation (61) yields equation (62).

(62) $$F = \frac{N_{out}}{N_{in}\mathrm{Loss}}$$

Numeric noise power available at the output is given by equation (57) and noise power available at the input of the network is simply kT_SB. Substituting these noise powers into equation (62) yields equation (63).

(63) $$F = \frac{kT_{Thevenin}B}{kT_sB \times Loss}$$

If the temperature of the two-port network is at 290°K, the source, two-port network, and load are in thermal equilibrium. The noise power at the output is equivalent to the noise power at the input and noise factor becomes:

(64) $$F = \frac{1}{Loss}$$

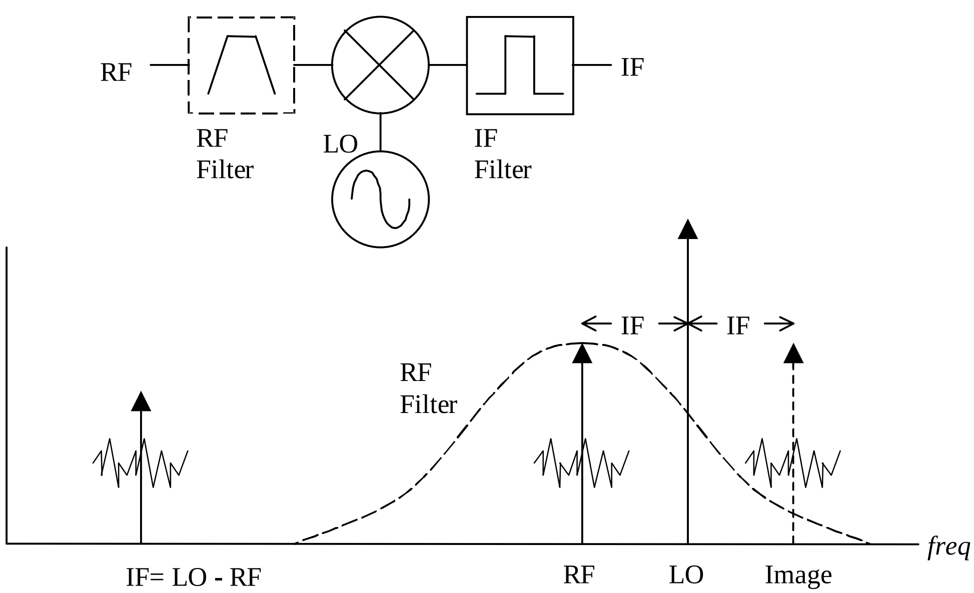
The loss (available gain G_a) in equation (64) is a number less than unity because S_{in} is greater than S_{out} for the two-port. Taking the reciprocal of the loss thus yields a result greater than unity. Converting equation (64) to dB results in a positive value. The noise figure for the lossy two-port network is shown by equation (65) to be equal to the loss in dB.

(65) $$NF = 10\log\left(\frac{1}{Loss}\right) = 10\log\left(\frac{S_{in}}{S_{out}}\right) = Loss_{dB}$$

Because the two-port network noise figure is equal to its loss, equation (51) predicts that adding a loss in front of a gain block degrades system noise figure dB for dB. Loss in front of a gain stage gives a total noise figure for the cascaded pair equal to the gain block's noise figure plus the loss of the component added. The system noise figure is optimized by avoiding or minimizing system losses. Added losses have less impact behind gain stages.

Mixer Noise Figure

Figure 28. Mixer Conversion Noise



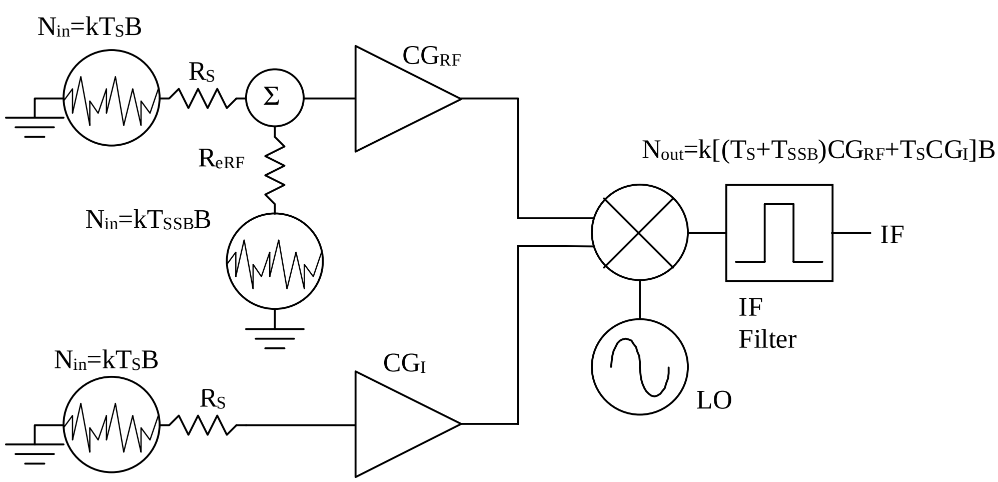
Noise power in the RF and image bands mixes with the LO and translates to the IF frequency whether RF and image signals are present or not as shown in Figure 28. A mixer's noise characterization is performed using one of two methods:

- Single-sideband (SSB)
- Double-sideband (DSB)

For single-sideband, output noise power is assumed to originate at the input from the RF frequency band. Double-sideband assumes that output noise power originates from both the RF and image frequency bands. Under either assumption, the mixer has both an image response and an RF response. Both responses usually have nearly the same conversion efficiency. When image and RF conversions are equivalent, the DSB noise temperature is half the SSB noise temperature; thus, the SSB noise figure is 3 dB greater than the DSB noise figure.

Figure 29 illustrates a single-sideband mixer noise model. Internal noise sources are removed and referenced to the input as shown. Conversion efficiency for the RF frequency CG_{RF} and the image frequency CG_I are removed and referenced at the input as well. The figure shows two inputs; however, only one input actually exists with two noise generating mechanisms. The total output noise is given by equation (66).

Figure 29. Single-Sideband Mixer Noise Model



(66) $$N_{out} = k\left[\left(T_S + T_{SSB}\right)CG_{RF} + T_S CG_I\right]B$$

Dividing by kB and solving for mixer SSB noise temperature T_{SSB} yields equation (67).

(67) $$T_{SSB} = \frac{T_{out} - T_S\left(CG_{RF} + CG_I\right)}{CG_{RF}}$$

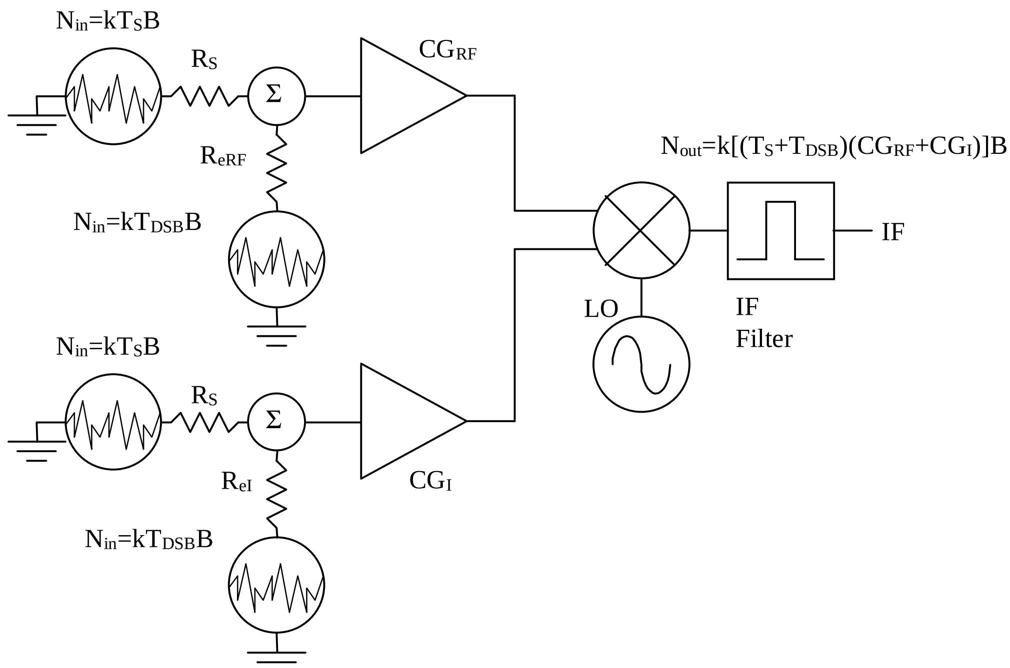
Figure 30 depicts a double-sideband mixer noise model. Internal noise sources are removed and referenced to the input as shown. Conversion efficiency for the RF frequency CG_{RF} and the image frequency CG_I are removed and referenced at the input as well. The total output noise is given by equation (68).

(68) $$N_{out} = k\left[\left(T_S + T_{DSB}\right)\left(CG_{RF} + CG_I\right)\right]B$$

Equations (66) and (68) are numerically equivalent. Dividing by kB and solving for mixer DSB noise temperature T_{DSB} yields equation (69).

(69) $$T_{DSB} = \frac{T_{out} - T_S(CG_{RF} + CG_I)}{CG_{RF} + CG_I}$$

Figure 30. Double-Sideband Mixer Noise Model



An RF filter in front of the mixer rejects noise in the image band, thus eliminating the need to consider SSB versus DSB noise response. In this case, the mixer SSB NF is used in system budget analysis.

Minimum Detectable Signal

A receiver's noise figure is an important parameter in determining the weakest signal the system can process. This translates directly into a maximum distance from the transmitter where communication is possible. The output noise power for a two-port network provides a method of quantifying and comparing receivers, as given by equation (30). The output noise power delivered to a matched load is a function of the two-port's noise temperature, noise equivalent bandwidth, noise factor, and gain factor. This is the system's noise floor referenced to the output. Equation (30) is segmented into four distinct parts, and noise temperature is set to a reference of 290°K to get equation (70).

(70) $$N_{out} = FkT_0BG_a = (F)(kT_0(1Hz))\left(\frac{B}{1Hz}\right)G_a$$

where:

F is the system's noise factor.
$kT_0 \times (1\text{ Hz})$ is the thermal noise power in a 1-Hz bandwidth at T_0
B is the noise equivalent bandwidth.
G_a is the two-port's numeric available gain.

Thermal noise power available from any loss in a 1-Hz noise equivalent bandwidth at a reference temperature of 290°K is given by equations (14) and (15) and repeated for convenience in equation (71).

(71) $$N_{1Hz} = kT_0(1Hz) = 4.002 \times 10^{-21}\text{Watt} = -174\text{dBm}$$

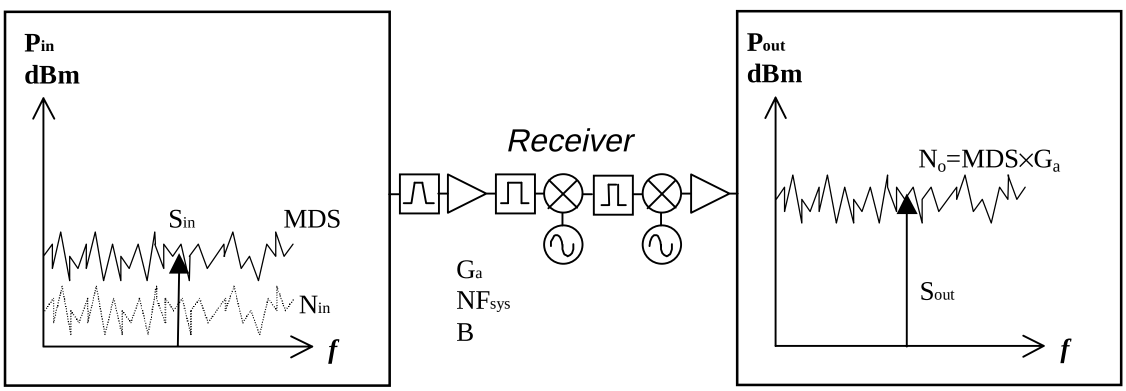
Equation (70) is converted to dB by taking 10 times the log of both sides of the equation thus obtaining noise output power in dBm.

(72) $$10\log(N_{out}) = 10\log\left[(kT_0(1Hz))\left(\frac{B}{1Hz}\right)(F)(G_a)\right]$$

$$= -174\text{dBm} + 10\log B + NF_{sys} + G_{dB}$$

Equation (72) is commensurate to a signal power S_{out} delivered to the output equal to the output noise floor. This equation defines the minimum detectable signal (MDS) and references the output. Figure 31 shows a dual-conversion receiver with associated noise and signal levels at the input and output. S_{out} is equivalent to the output noise power N_{out} ($MDS \times G_a$) as shown.

Figure 31. Dual-Conversion Receiver With Associated Noise and Signal Levels



The input referenced noise floor is calculated by subtracting the system gain in dB from the output noise power, thus giving equation (73).

(73) $$MDS = -174\text{dBm} + 10\log B + NF_{sys}$$

The MDS equation determines the input signal level required to deliver an output signal to a load equivalent to the output noise floor. The noise floor is directly proportional to bandwidth as the equation shows. Thus, to lower a receiver system's noise floor, the equivalent noise bandwidth needs to be as narrow as possible without filtering out portions of the desired signal. Higher data rates require more bandwidth than lower data rates. High data rate systems constrain the narrowness of the equivalent noise bandwidth. This constraint, in turn, limits the receiver's minimum detectable signal.

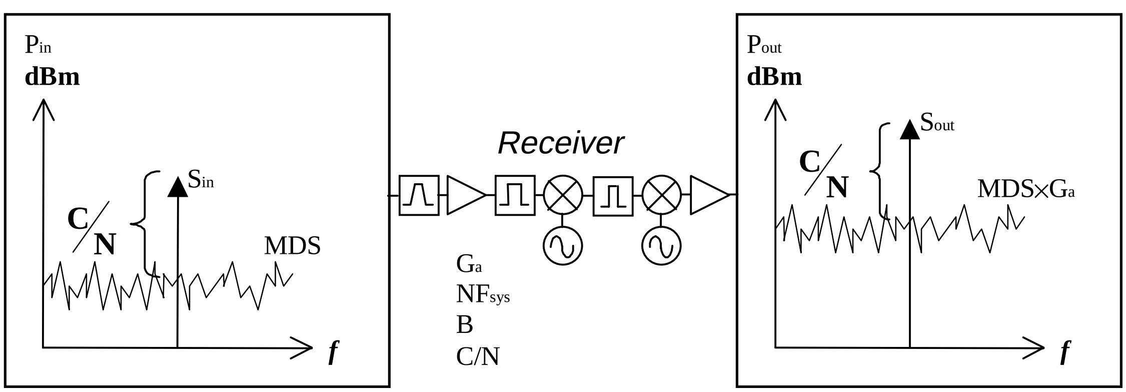
A tradeoff of data rate and minimum detectable signal always exists for a given modulation scheme. The system's noise figure impacts the MDS equation dB for dB. The noise equivalent bandwidth and system noise figure translate directly to distance from the transmitter for reliable communication. The MDS equation calculates a signal power level at the input that takes into account noise added by the system. MDS therefore represents a signal level presented to the input of a receiver that would deliver an output signal to the load equivalent to the noise power delivered to the load. The numeric input signal level, S_{in}, shown by Figure 31, is the MDS level at the input. The value of S_{in} is given by the product of the input noise power N_{in} and the two-port's noise factor. The actual noise power available from the source N_{in} is shown in the figure by the dashed noise plot and given by kT_0B. Noise available from the source, in the receiver noise equivalent bandwidth, is converted to dBm by equation (74).

(74) $$10\log(N_{in}) = -174\text{dBm} + 10\log B$$

Signal-to-Noise Ratio

The MDS equation calculates an output signal power equal to the output noise floor. Most receivers require a desired signal above the noise floor to recover information. In other words, a specific signal-to-noise ratio is required to recover the information with a specific quality level. The required signal-to-noise ratio is also referred to as carrier-to-noise ratio. Figure 32 shows a dual-conversion receiver and the carrier-to-noise ratio (C/N) at the input and the output.

Figure 32. Dual-Conversion Receiver and Carrier-to-Noise Ratio



MDS references source noise and noise added by the receiver to its input terminals. This is not the "true" noise floor but a "fictitious" noise floor due to noise added by the receiver. The "true" noise floor is simply the source noise given by equation (74) for N_{in} (kT_SB) when the receiver is terminated by a passive termination at its input. (Figure 31 shows N_{in} and MDS at the input of a receiver.)

The carrier-to-noise level is the same at both the input and the output when MDS is used as the noise reference at the input. The input signal S_{in} must be above the MDS level, as shown in Figure 32. This means that the output signal S_{out} is above the output noise floor by the same amount. The “true” signal-to-noise ratio at the input is greater than the signal-to-noise ratio at the output when the source noise is used as a reference since the receiver system adds noise. The required signal-to-noise ratio depends on several system parameters. The modulation scheme, IF filters’ group delay distortion, detector linearity, distortion, and several other parameters contribute to the required signal-to-noise ratio.

The quality of received information also plays a large role in the carrier-to-noise required. In a digital receiver, the bit-error rate (BER), or the probability P(e) that any received bit is an error, is the primary quality measure. BER is defined as the number of bit errors divided by the number of bits transmitted. Jitter, which results from noise in the receiver, degrades BER performance. Detector linearity, frequency response of the filter, bandwidth, and other parameters contribute to the errors.

Figure 33. Probability of Error Versus Carrier-to-Noise Ratio

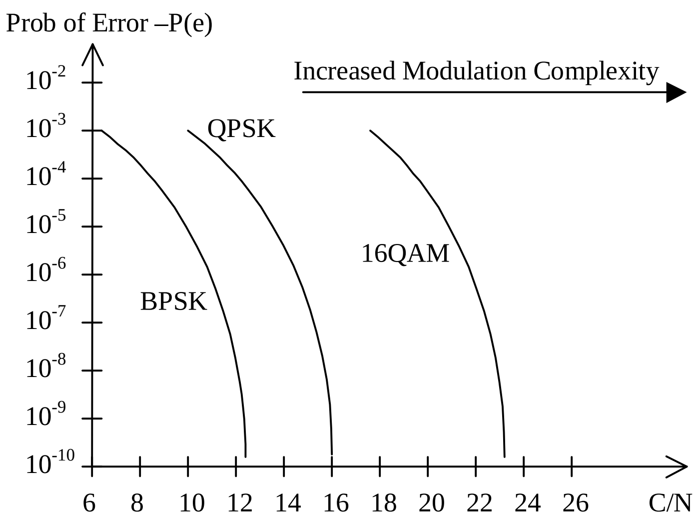


Figure 33 shows the probability of error versus carrier-to-noise ratio for various modulation schemes. BER has a nonlinear relationship to the receiver’s MDS. The higher the desired signal is above the noise floor, the lower the BER (P(e)). As signal-to-noise ratio decreases, the probability of error increases dramatically. BER changes by several orders of magnitude with small changes in the signal-to-noise ratio. Consequently, the message reliability decreases with a decreased signal level.

Modulation schemes with increased complexity require higher signal-to-noise ratios for an acceptable bit error rate. Typically, the higher the system data rate the more complex the modulation scheme. If bits are transmitted individually, an increased data rate requires an increase in noise bandwidth. The increased bandwidth degrades the noise floor given by the MDS equation. Modulation schemes that send symbols instead of bits reclaim some of the lost receiver noise bandwidth.

Sensitivity

Sensitivity is defined as the signal level required for a particular quality of received information. For digital radios, quality is measured by the bit error rate. A specific signal-to-noise ratio is required for a given bit error rate. Sensitivity is the absolute power level that gives the required signal to noise ratio. Sensitivity is computed based on the MDS and required carrier-to-noise ratio, as depicted in Figure 32. Mathematically, sensitivity referenced to the input is defined as the sum of the MDS and required output signal-to-noise ratio given by equation (75).

(75) $$\text{Sensitivity}_{dBm} = \text{MDS}_{dBm} + C/N$$

where:

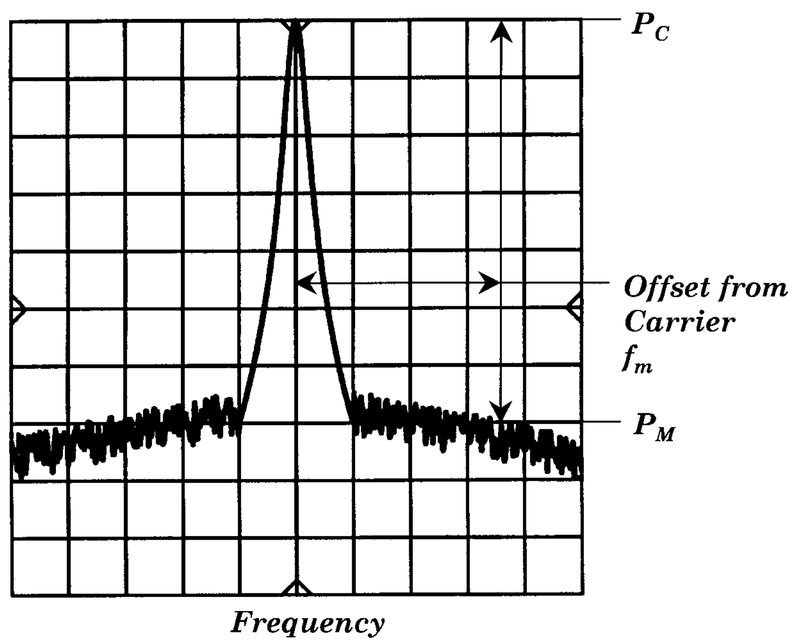
MDS is the minimum detectable signal.

C/N is the required output signal-to-noise ratio for a specific quality of received information.

Phase Noise

Phase noise describes an oscillator’s short-term random frequency fluctuations. The modulating signal is noise that yields sidebands forming a noise spectrum in the oscillator circuit. Noise sidebands appear on both sides of the carrier frequency. Noise sidebands are typically specified in terms of dBc/Hz (amplitude referenced to a 1-Hz bandwidth relative to the carrier) at a given offset frequency from the carrier frequency. Typically, one sideband is considered when specifying phase noise, thus giving single sideband performance. Figure 34 shows a double-sideband phase noise plot of a signal source.

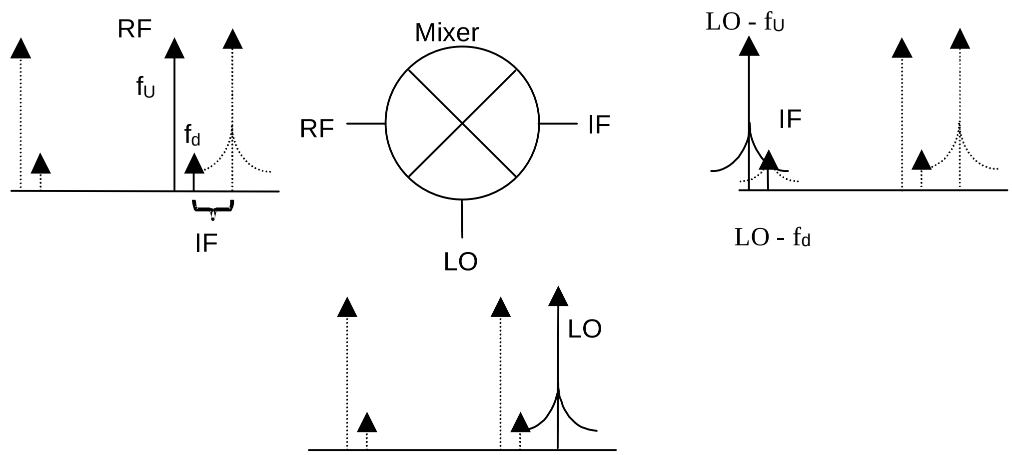
Figure 34. Double-Sideband Phase Noise Plot of a Signal Source



Phase noise is injected into the mixer LO port with the oscillator signal shown in Figure 35. If perfect sinusoidal signals are input to the RF port (no phase noise), the LO signal and its phase noise mixes with the "clean" RF signals. This produces IF signals containing phase noise. If a small desired signal f_d and a large undesired signal f_U nearby in frequency are input to the RF port, the phase noise on the larger conversion signal may mask the smaller desired signal as shown in Figure 35. This would hinder reception of the desired signal. Thus, low phase noise is crucial for oscillators in receiver systems.

Phase noise also adds to the RMS phase error during the detection of digitally modulated signals.

Figure 35. Conversion Limitations Due to Phase Noise



Receiver Nonlinear Performance

Receiver systems process very small desired signals and reject very large undesired signals. The weak signals are not processed if the large undesired signals interfere with reception. Increasing the desired signal's amplitude is one method to effectively counter such interference. Raising the output power of the transmitter can increase signal amplitude.

Raising transmitter output power can require power increase in several geographic locations to overcome large undesired signals in the service area. Therefore, this option is not always practical. Instead of concentrating on the transmitter, improvement in receiver performance and rejection of interfering signals may solve most of the message reliability problems. Clearly, the interfering signal must be decreased such that the distortion amplitude is somewhat below the weak signal amplitude.

Gain Compression

In the linear range of a two-port network, a change in the input signal produces a change in the same direction dB for dB in the output signal. Numeric available gain, a small-signal gain value, is defined as the ratio of available output power to available input power. Mathematically, numeric available gain is defined by equation (76).

$$G_a = \frac{P_{out}}{P_{in}} \quad (76)$$

where:

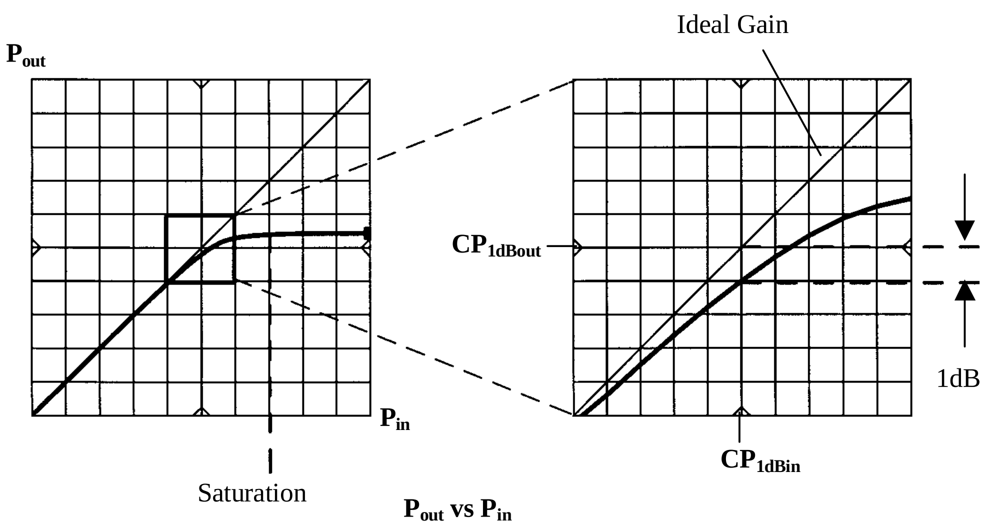
P_{out} is the power available to the load in Watts.
P_{in} is the power available from the source in Watts.

Numeric gain is converted to dB using equation (77).

$$G_{dB} = 10\log\left(\frac{P_{out}}{P_{in}}\right) \quad (77)$$

With an ideal two-port, the ratio of equation (77) is constant for all input power values. Figure 36 shows the ideal two-port P_{out} versus P_{in} characteristic.

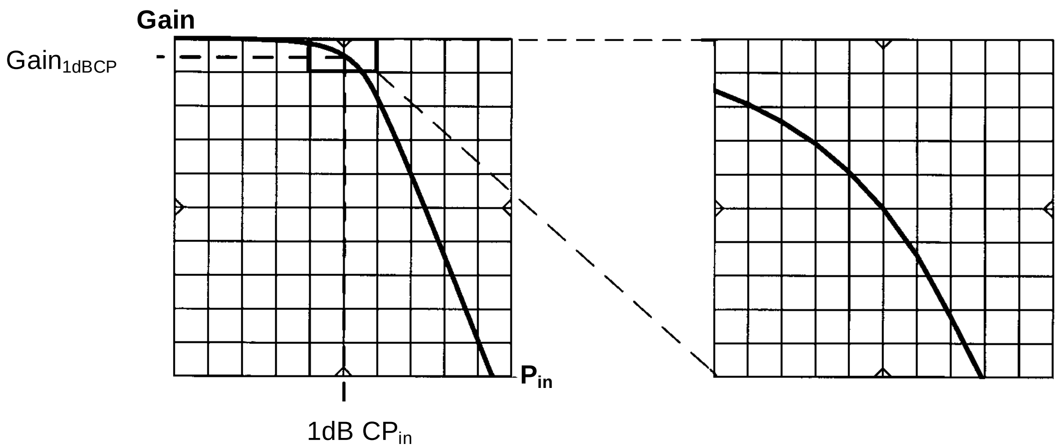
Figure 36. Ideal Two-Port P_{out} Versus P_{in} Characteristic



As the input signal increases for a practical two-port network, the output signal increases linearly until distortion product energy substantially combines with the fundamental output power. The result is that distortion products, along with fundamental, are observed at the output. The linear output power change versus an input power change no longer exists. The ratio of equation (76) is no longer constant. At some input signal level, known as the 1-dB compression point (1-dB CP), the output power is 1 dB below the output power of an ideal two-port. As the input signal continues to increase, higher order harmonics also contribute and are observed at the output. The output power versus input power continues to deviate from the ideal two-port network. Eventually, the input signal increases to a point where fundamental frequency output power reaches a maximum value. This point is the saturation point, as shown in Figure 36.

Figure 37 plots two-port gain versus input power level. Low input power values provide a constant ratio in equation (76) and the small signal gain. As input signal level increases, the gain drops (or loss increases). At a particular input power level, gain drops 1 dB below small signal gain of equation (77).

Figure 37. Two-Port Gain Versus Input Power Level



Once the two-port is compressed, gain is decreased for all signals being processed through the circuit. A strong signal (either desired or undesired) input into a receiver can cause compression. Consider a weak desired signal along with a strong undesired signal input into the receiver. The strong undesired signal input into the receiver causes compression. If the desired signal is at the sensitivity level, a decrease in gain reduces the carrier-to-noise level, which in turn increases the bit error rate. Message reliability deteriorates.

The 1-dB gain compression point is defined as the amount of power required to cause a gain drop of 1 dB due to circuit nonlinearity. The 1-dB gain compression point is referenced either to the input or output. A 1-dB gain compression referenced to the input (CP_{1dBin}) indicates the amount of power incident at the input of the two-port that reduces the small signal gain of equation (77) by 1 dB. Alternately, when referenced to the output (CP_{1dBout}), it is defined as the amount of fundamental power delivered to the output load when the two-port is compressed by 1 dB. The 1-dB gain compression point referenced to the output 1-dB gain compression point (CP_{1dBout}) is determined from gain compression referenced to the input CP_{1dBin} by subtracting the small-signal gain and adding 1 dB mathematically

(78) $$CP_{1dBout} = CP_{1dBin} + \text{Gain} - 1\,\text{dB}$$

If the 1-dB gain compression point is referenced to the output, rearranging equation (78) yields 1-dB compression point referenced to the input CP_{1dBin}.

(79) $$CP_{1dBin} = CP_{1dBout} - \text{Gain} + 1\,\text{dB}$$

Intermodulation Distortion

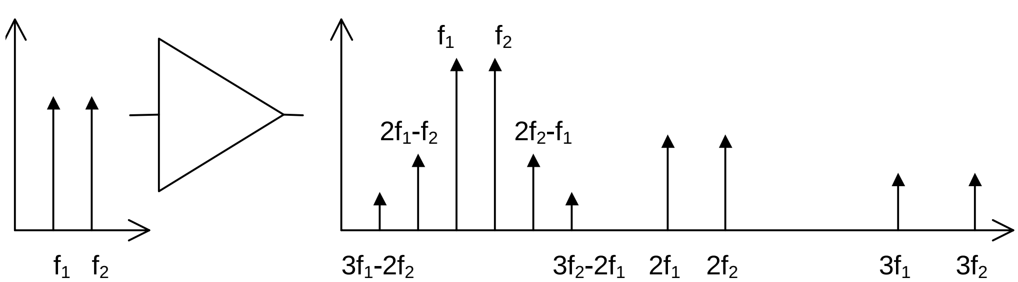
Receiver circuits generate other distortion products (terms). Consider two RF signals, f_1 and f_2, connected to the input of a two-port network, as shown in Figure 38. The frequency spacing between the two example signals f_1 and f_2 is 25 kHz. The frequency domain plot depicts two input signals. First, assume the two-port is perfectly linear. Both signals are amplified and delivered to the output load, as shown in Figure 38. Changing both signals input power yields a corresponding output power change. For example, increasing input power by 1 dB results in a corresponding increase of 1-dB fundamental output power. The smaller spectral lines do not exist for the perfectly linear two-port.

Next consider a nonlinear two-port network for Figure 38. This two-port outputs two signals, f_1 and f_2, as discussed. Additional signals known as distortion products are generated internally and delivered to the load. A sum term $f_1 + f_2$, a difference term $f_2 - f_1$, and the second harmonics $2f_1$ and $2f_2$ are output. (These terms are not illustrated in the figure.) Additionally, mixing products are generated at easily determined frequencies given by equation 80.

(80) $$IMD = \pm m f_1 \mp n f_2$$

Third-order IMD products occur at frequency locations $2f_2$-f_1 and $2f_1$-f_2 given by equation (80). Thus, the second harmonic of f_1 mixes with fundamental f_2 and the second harmonic of f_2 mixes with fundamental f_1. These distortion products are equally spaced from the two fundamental frequency terms. For this example, the fundamentals are spaced 25 kHz apart, thus, a third-order distortion product appears 25 kHz above the two original frequencies f_1 and f_2 and the other appears 25 kHz below the two original frequencies. The second harmonic of f_2 mixes with the fundamental f_1. These distortion products are equally spaced from the two fundamental frequency terms. For this example, the fundamentals are spaced 25 kHz apart, thus, a third-order distortion product appears 25 kHz above the two original frequencies f_1 and f_2 and the other appears 25 kHz below the two original frequencies. These mixing products described by $2f_2$-f_1 and $2f_1$-f_2 are known as third-order intermodulation distortion products depicted by the smaller spectral lines in Figure 38. Changing f_1 and f_2 power levels by 1 dB changes both third-order IMD products by 3 dB.

Figure 38. Receiver Circuits and Intermodulation Distortion



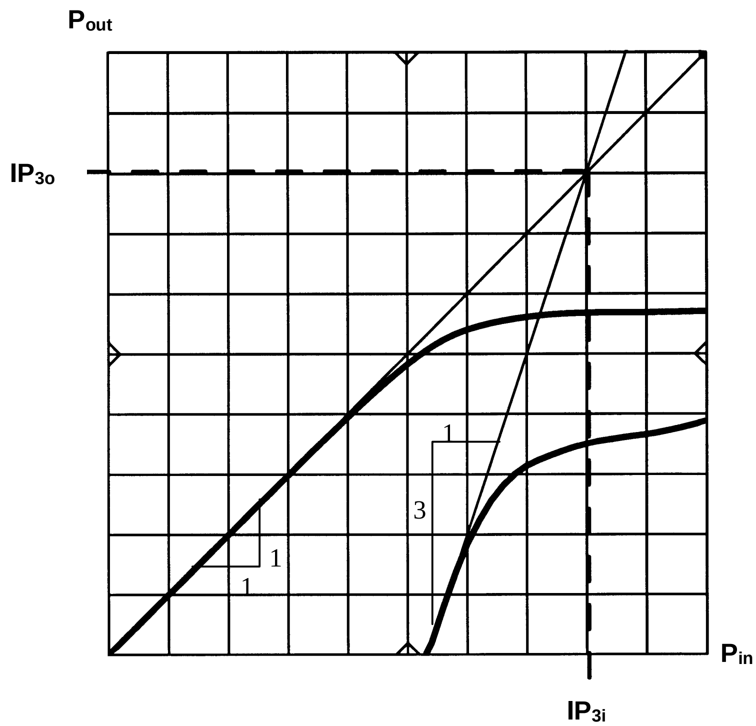
The power difference between the fundamental frequency output power and distortion product power is known as the rejection ratio. Third-order intermodulation rejection ratio decreases by 2 dB for every 1-dB input power increase. Theoretically, a continual input power level increase eventually leads to distortion product power equal to fundamental frequency power delivered to the load. This power level of equality is known as the two-port device's third-order intercept point (TOI or IP3). A network such as an attenuator that adds no distortion has an infinite intercept point.

Figure 39 plots output power, P_{out}, for one fundamental frequency signal versus input power, P_{in}. Fundamental output power versus input power has a linear characteristic with unity slope as shown. The figure plots one third-order intermodulation distortion product along with the fundamental. The third-order intermodulation distortion product has a linear characteristic with a slope of three. Because the distortion product slope is steeper than the fundamental power slope, the third-order distortion characteristic eventually intersects the fundamental curve. This intersection is known as the two-port's third-order intercept point. Intercept point is referenced at the two-port input or output. Third-order intercept point referenced to the input is denoted as IP_{3i} and shown in Figure 39. Third-order intercept point referenced to the output is denoted as IP_{3o}. The output intercept point is equivalent to the input intercept point plus two-port small-signal gain in dB as shown by equation (81).

(81) $$IP_{3o} = IP_{3i} + Gain$$

A practical two-port network cannot be driven with enough power to reach the theoretical intercept point. Compression occurs before reaching the point of intersection, as shown in Figure 39. Once the device starts compressing, fundamental frequency and distortion product output power limit with increased input power. The 1-dB compression point is typically around 10 dB below the third-order intercept point. Intercept point measurements must be performed at power levels where the device is not compressed.

Figure 39. Output Power P_{out} for One Fundamental Frequency Signal Versus Input Power P_{in}



The intercept point permits a full description of the third-order intermodulation distortion as long as the two-port is not compressed. Simple linear relationships allow determination of distortion product power for a given fundamental input or output power level.

Figure 40. Fundamental P_o and Third-Order Distortion P_d Output Powers Versus Input Power P_{in}

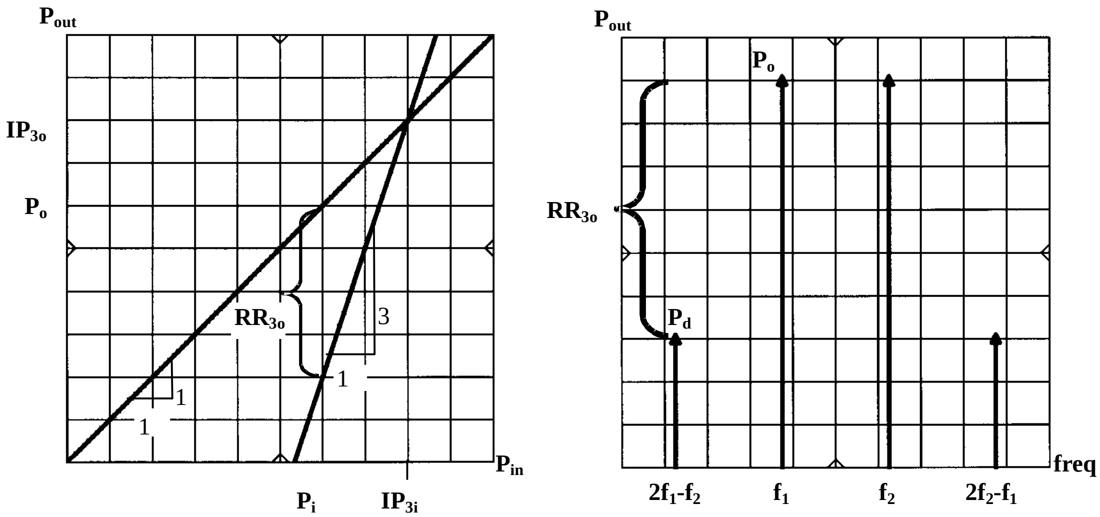


Figure 40 illustrates the linear relationship of fundamental P_o and third-order distortion P_d output powers versus input power P_{in}. The output frequency spectrum at a fixed input power level is included. The third-order rejection ratio is easily measured using a spectrum analyzer at a fixed input power level. Rejection ratio is the difference in dB between fundamental P_o and distortion P_d output powers. The rejection ratio measurement defines one point on the third-order distortion characteristic. A line with a slope of three is constructed on the same rectangular grid as the fundamental output power versus input power characteristic. Third-order rejection ratio RR_{3o}, referenced to the output, is defined on the output power versus input power graph as shown. For a given input power P_{in}, rejection ratio RR_{3o} is graphically determined as the difference in dB between the fundamental output power line and the third-order distortion product line.

The rejection ratio for a given power level is needed to determine if undesired signals will interfere with a receiver in a particular environment. The rejection ratio is shown on the P_{out} versus P_{in} graph as well as the frequency domain plot. The output third-order intercept point is obtained by measuring the output power and rejection ratio and using equation (82).

(82) $$IP_{3o} = P_o + \frac{RR_{3o}}{2}$$

where:

Po is the output power of one fundamental.
RR_{3o} is the third-order rejection ratio at P_o.

Third-order rejection ratio RR_{3o} is simply the dB difference between the measured output power of one of the fundamental tones and one of the distortion products, mathematically given by equation (83).

(83) $$RR_{3o} = P_o - P_d$$

Substituting equation (83) into equation (82) yields the third-order output intercept point of the two-port.

(84) $$IP_{3o} = P_o + \frac{P_o - P_p}{2} = 1.5P_o - .5P_p$$

Measuring fundamental output power P_o and distortion power P_d determines the third-order intercept point referenced to the output.

Third-order input intercept point IP_{3i} is determined by rearranging equation (81) to yield equation (85).

(85) $$IP_{3i} = IP_{3o} - \text{Gain}$$

Given the third-order output intercept point, the distortion product power level is easily determined for other input power levels. Rearranging equation (84) gives the distortion product power at a different output power level.

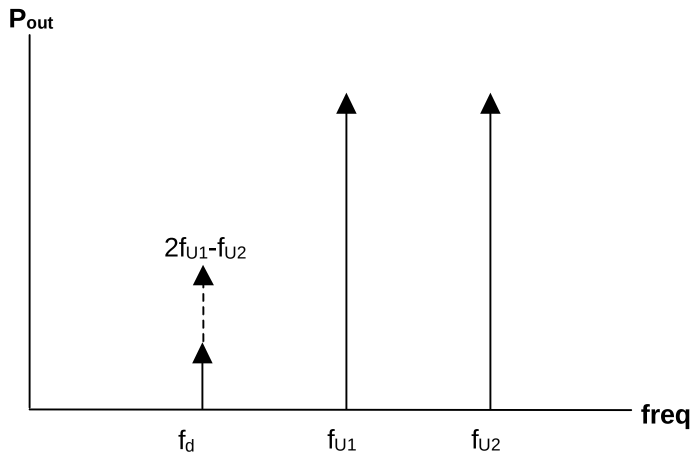
(86) $$P_p = 3P_o - 2IP_{3o}$$

Third-order intercept point is an important performance parameter in an environment with signals within the receiver's RF passband. A whole range of frequencies pass unimpeded through the RF passband for receiving selected channels within that passband. These adjacent and close-in signals process through nonlinear circuits in the receiver's front-end, such as the amplifier and mixer circuits. Because close channels pass unimpeded, they create third-order distortion products within the RF front end, including the first mixer. These distortion products occur at the same frequency as the desired signal. If the distortion product is at a power level slightly less than or at a power anywhere above the desired signal, it interferes with message reception.

Figure 41 shows a typical output spectrum of a two-port receiver network. First, consider a perfectly linear two-port receiver. The frequency domain plot shows the desired signal f_d and two undesired signals f_{U1} and f_{U2}. The closest undesired signal f_{U1} is one channel spacing away from the desired signal, in this case 25 kHz. The second undesired signal f_{U2} is two channel spacings away from the desired signal, in this example 50 kHz. (The desired signal is shown as the short solid line.) Interference is not possible from third-order distortion products in the perfectly linear receiver case. Messages are reliable if the desired signal is at or above the sensitivity level.

Now assume the receiver is not perfectly linear. The two undesired signals f_{U1} and f_{U2} cause a third-order intermodulation distortion product coincident with the desired signal. (The third-order intermodulation distortion product is the dashed line.) The distortion product power is greater than desired signal power, thus prohibiting message reception. Once the distortion product is coincident with the desired signal, filtering cannot ensure clear reception. Therefore, the intercept point must be high enough to provide a third-order rejection ratio capable of protecting against the largest interfering signals incident at the receiver input terminals.

Figure 41. Typical Output Spectrum of a Two-Port Receiver Network

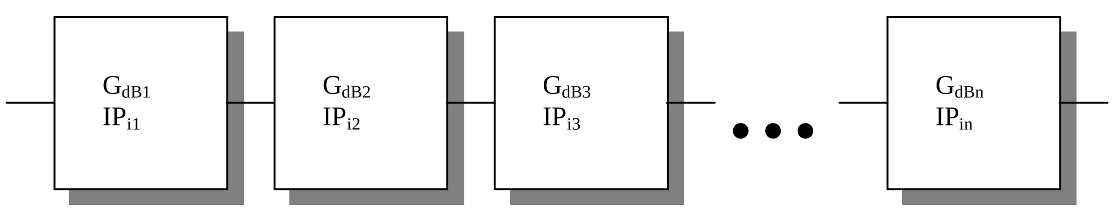


Intermodulation distortion products are also created by signals one channel and two channel spacings below the desired signal. Additionally, third-order intermodulation distortion is caused by undesired signals located 2 channel and 4 channel spacings away from the desired signal. Interference from undesired signals located 3 channel and 6 channel spacings away is also possible. Other combinations throughout the receiver RF passband produce third-order distortion products coincident with the desired signal.

A receiver contains several cascaded networks or blocks, each of which performs a particular function to process the desired signal or reject the undesired signals. Each network may or may not add distortion during signal processing. The third-order intercept point is defined for individual blocks in the receiver chain. Having third-order intermodulation distortion information for each network allows the total system intercept point to be determined. The system intercept point is used to predict message reliability in a particular environment.

Figure 42 portrays a receiver containing several cascaded networks with known gains and third-order input intercept points. Each network in the chain has a particular impact on total system distortion, even if the individual circuit does not add distortion of its own.

Figure 42. Cascaded Receiver Networks



Total Gain; G_{dBT}
Total Input Intercept Point; IP_{3iT}

The system intercept point is computed with an equation similar to the cascaded noise factor equation (equation (51)). Equation (87) uses numeric gains and input intercept points. Total numeric input intercept point ip_{iT} for any undesired response is determined by the reciprocal of equation (87).

(87) $$\frac{1}{ip_{iT}} = \left[\left(\frac{1}{ip_{i1}}\right)^q + \left(\frac{g_1}{ip_{i2}}\right)^q + \left(\frac{g_1 g_2}{ip_{i3}}\right)^q + \cdots + \left(\frac{g_1 g_2 \cdots g_{n-1}}{ip_{in}}\right)^q\right]$$

where:

(88) $$q = \frac{m-1}{2}$$

m is the slope of the undesired response (*m* = 3 for third order).

Equations (89) and (90) convert intercept points and gains from dB to numeric form.

(89) $$ip_{i1} = 10^{IP_{i1}/10} \cdots ip_{in} = 10^{IP_{in}/10}$$

(90) $$g_1 = 10^{G_1/10} \cdots g_{n-1} = 10^{G_{n-1}/10}$$

The cascaded intercept point equation does not use the gain of the last stage. System numeric input intercept point is converted to dB using equation (91).

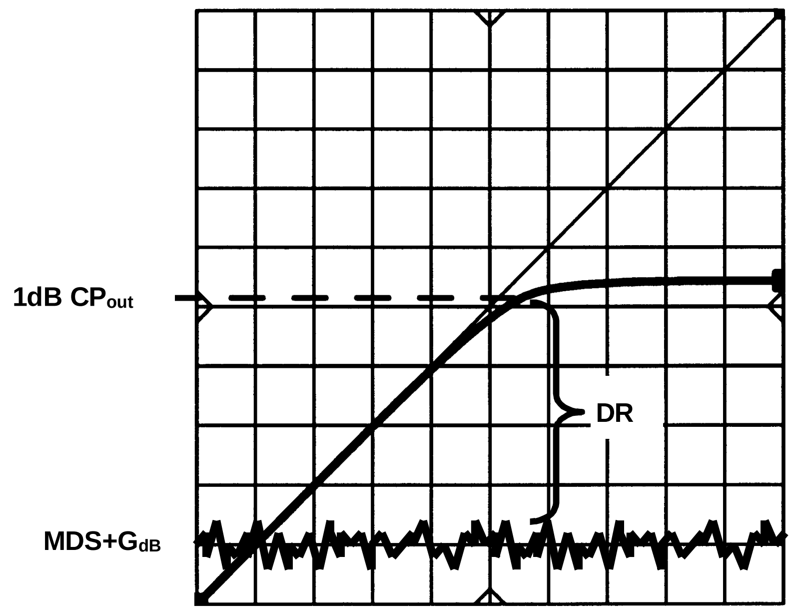
(91) $$IP_{3iT} = 10\log(ip_{3iT})$$

Total intercept point is given in dBm if the numeric power is referenced to 1 mW. Loss in front of the system improves input intercept point dB for dB. This is a direct conflict for noise figure, as determined in the section, *Receiver Building Blocks*, where adding loss in front of a receiver system degrades noise figure dB for dB of loss added. Increasing gain in front of the system degrades the entire system's intercept point but improves noise figure. The design process requires a trade-off between noise figure and intercept point.

Dynamic Range

Dynamic range is the useful signal level range the receiver can process with a particular information quality. Bit error rate measures quality, as is the case with digital radios. Dynamic range may be defined as the difference in power level between the 1-dB compression point and the system noise floor. The system noise floor referenced to the input is given in the section, *Receiver Building Blocks*, as the minimum detectable signal (MDS). Adding the small signal system gain to the MDS yields noise floor referenced to the output. The 1-dB compression point is referenced at the receiver input or output given by equations (78) and (79). Figure 43 illustrates dynamic range based on MDS and 1-dB compression point referenced to the output.

Figure 43. Dynamic Range



Dynamic range is calculated by subtracting the MDS from the 1-dB compression point referenced to the input. In equation form:

(92) $$DR = CP_{1dBin} - MDS_{dBm} = CP_{1dBin} + 174\ dBm - 10\log(B) - NF_{sys}$$

The 1-dB compression point CP_{1dBin} is given in dBm.

Receiver dynamic range is sometimes modified from the above definition using sensitivity instead of MDS on the low end. This definition gives a value less than the definition of equation (92) by an amount equal to the required carrier-to-noise ratio. Mathematically,

(93) $$DR = CP_{1dBin} - Sensitivity = CP_{1dBin} + 174\ dBm - 10\log(B) - NF_{sys} - C/N$$

The top end of the dynamic range definition may also be modified for receiver systems. Instead of using the 1-dB compression point, the highest power that the receiver can handle and still recover information at a particular quality level is used. The dynamic range is then defined as the difference between this power and the sensitivity power level.

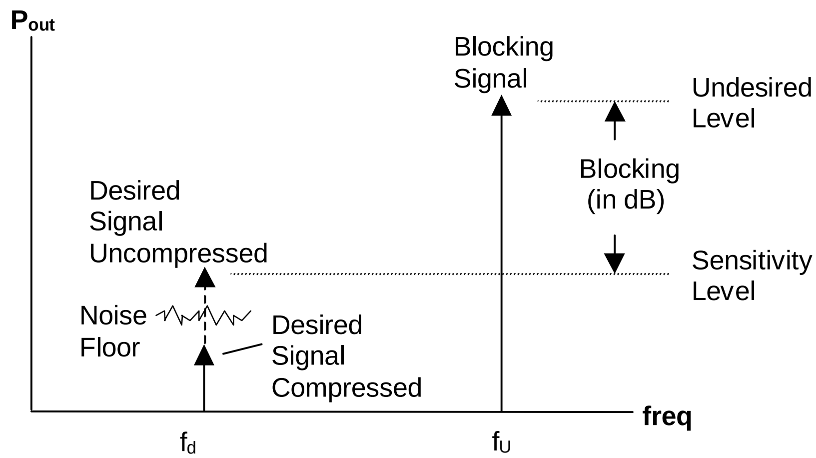
Dynamic range may also be defined as the difference between the largest signal the receiver can reject while receiving a smaller signal at sensitivity level. This description is referred to as receiver blocking.

Blocking

Large undesired signals within the receiver's RF passband cause gain compression of certain stages. Gain compression reduces the signal power delivered to the load for all signals including the desired signal. If the system gain drops, the total system noise figure may increase depending on which stage has been compressed. Equation (51) predicts system noise figure partially as a function of distributed gains. If noise figure degrades, the MDS and sensitivity levels degrade. The noise figure of the system may not degrade significantly, but if gain compression is realized, the amount of desired signal delivered to the load is reduced. Both effects degrade signal-to-noise ratio. Consider a receiver with a desired signal at the receivers' sensitivity level. If a large undesired signal is introduced within the RF passband, the receiver gain drops due to compression. The gain drop causes a reduction in the desired signal, thus lowering signal-to-noise ratio. This reduction, in turn, causes the desired signal to drop below the sensitivity level. The sensitivity level requires a particular signal-to-noise ratio to recover information at a specified quality level. The reduction in signal-to-noise ratio thus causes a loss in recovered information.

Digital radios use a specified bit error rate performance. *Blocking* is defined as the amount of power required for receiver sensitivity degradation from compression effects. Blocking is usually characterized with an undesired signal located 1 MHz away from the desired signal. The difference between the undesired signal power and the sensitivity level is known as receiver blocking and is expressed in dB.

Figure 44. Receiver Blocking



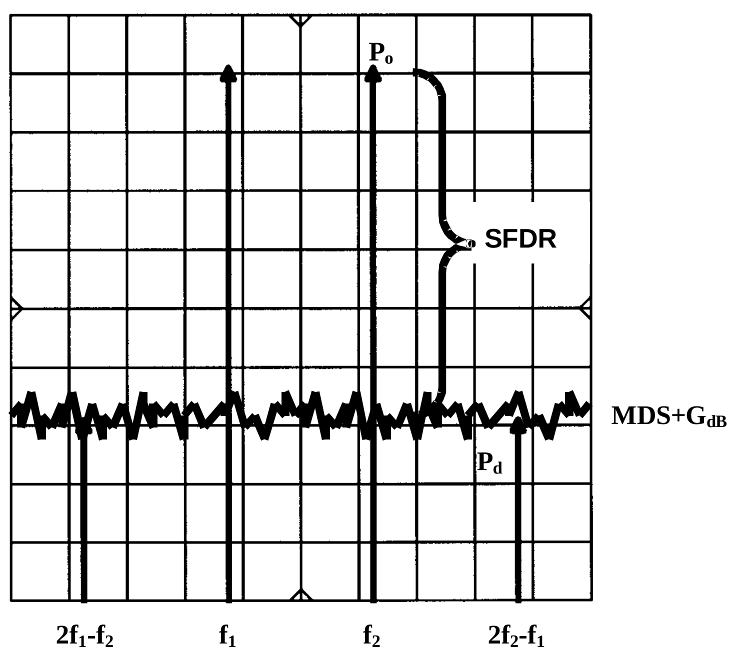
Spur-Free Dynamic Range

Spur-free dynamic range is similar to third-order intermodulation distortion rejection ratio. The third-order intermodulation rejection ratio is defined as the power difference between the fundamental and distortion product output powers, given by equation (83). The rejection ratio depends on a specific output power. A change in fundamental power produces a change in rejection ratio. A 1-dB change in fundamental power gives a 3-dB change in third-order distortion product power. Because the fundamental has a slope of unity and the distortion product has a slope of 3, rejection ratio exhibits a slope of 2 versus fundamental output power. Thus, a 1-dB change in fundamental power changes rejection ratio by 2 dB.

The spur-free dynamic range is the difference between the fundamental power and the noise power when the distortion products are equal to the noise power. Figure 45 clarifies the spur-free dynamic range. Said another way, at a specific fundamental power level, the distortion product power is equal to the noise power. The spur-free dynamic range is defined mathematically by equation (94).

$$\text{(94)} \quad \mathrm{SFDR} = \frac{2}{3}\left[IP_{3o} - \mathrm{MDS} - G_{sys}\right] = \frac{2}{3}\left[IP_{3o} + 174\ \mathrm{dB} - 10\log(B) - NF_{sys} - G_{sys}\right]$$

Figure 45. Spur-Free Dynamic Range



Undesired Spurious Responses

Mixer circuits translate RF signals to some other frequency known as the intermediate frequency (IF). Mixing the RF signal with an LO signal and selecting the sum term (LO + RF) is known as up-conversion. If the mixer translates the RF signal to a lower frequency, the process is known as down-conversion. The down-conversion process is the most common for receivers and utilizes the difference term (LO – RF or RF – LO).

Two cases for down-conversion are possible. High-side injection has the LO frequency above the desired RF signal. Low-side injection has the LO frequency below the desired RF signal. The intermediate frequency is the difference frequency between the RF and LO frequencies. In equation form:

(95) $$f_{IF} = \pm f_{RF} \mp f_{LO}$$

where:

f_{IF} is the intermediate frequency output at the IF port.
f_{RF} is any RF signal available at the RF port.
f_{LO} is the local oscillator signal injected into the mixer LO port.

Because the mixer circuit is a non-linear device, other undesired mixing terms, called spurious responses, are possible. Spurious responses are caused by undesired signals reaching the mixer's RF input port and producing a response at the IF frequency. The undesired signal at the RF port does not have to fall within the RF passband to cause interference. When a receiver is subjected to spurious response signals, interference with the desired signal is possible. Mixer spurious responses are easily defined by equation (96).

(96) $$f_{IF} = mf_{RF} + nf_{LO}$$

where:

$m = \pm 1, \pm 2, \pm 3 \cdots$ and $n = \pm 1, \pm 2, \pm 3 \cdots$

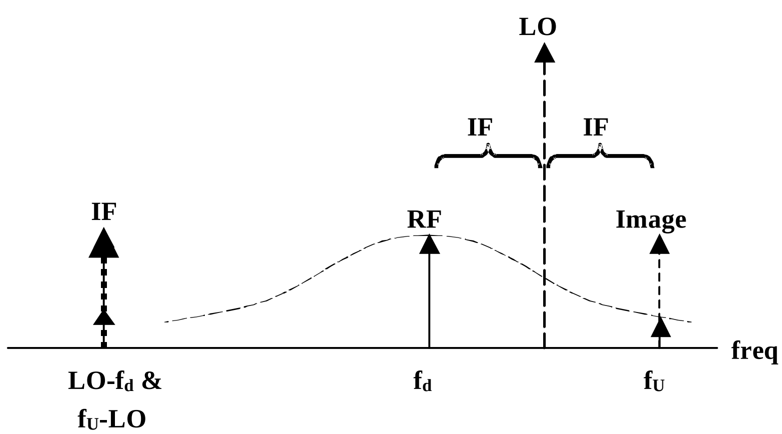
The integers *m* and *n* are harmonics of the RF and LO frequencies respectively.
The interfering signals have equal amplitudes in Figure 46 through Figure 48.

The most common spurious responses are the image, half-IF, and LO spurs. The RF, LO and IF frequencies are carefully selected to avoid spurious responses. Filters protect against RF signals outside the RF passband that cause an IF response. Some mixer designs protect against certain spurious responses. Image reject mixers protect against the image frequency. Balanced mixers reject some if not all spurious responses where *m* or *n* is even. LO filters also protect against certain types of spurious responses.

Image Response

Image is a spurious response with integers *m* and *n* equal to unity. Figure 46 shows a high-side injection example with the desired RF frequency f_d below the LO injection frequency. (The LO signal is injected into the LO port and shown by the dashed spectral line. If the f_U frequency had been the desired signal with low-side injection, then f_d would be the image frequency.) Solid spectral lines show the RF and image signals, as presented to the RF port. The IF frequency is available from the IF port shown with thick dotted spectral line. The desired RF signal f_d mixes with the LO signal to produce a difference term known as the IF frequency. The undesired signal called the image frequency f_U also mixes with the LO signal to produce an IF frequency. The image frequency interferes with the reception of the desired RF signal f_d. For protection from this spur, the image frequency must be filtered before it reaches the mixer. This filtering is usually performed by an image-reject RF filter prior to the mixer. Alternately, an image reject mixer protects from the image frequency as well.

Figure 46. High-Side Injection Example



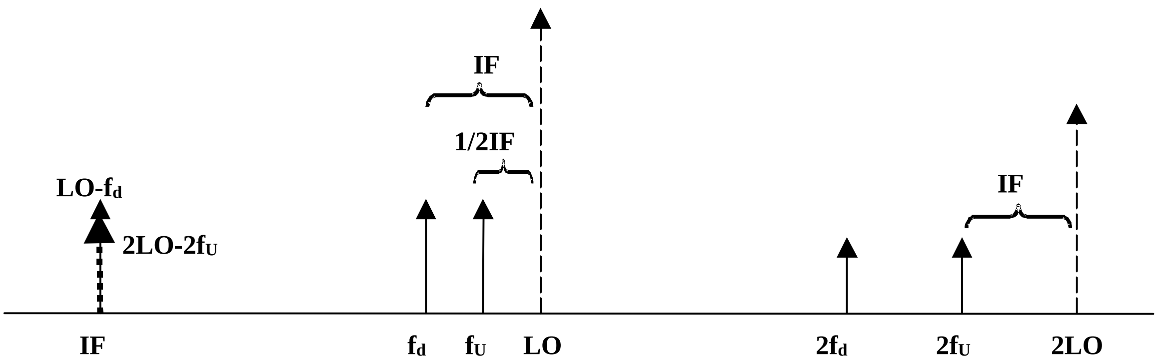
Half-IF Spurious Response

The half-IF spur is a fourth-order response where both the *m* and *n* integers equal two. The half-IF response is an interfering signal occurring halfway between the desired RF signal and the LO signal, as shown in Figure 47. The LO signal and its second harmonic, 2LO, is shown with a dashed spectral line. The RF and half-IF signals and the second harmonics of each are shown by the solid spectral lines.

Figure 47 shows a high-side injection example with the desired RF frequency f_d below the LO injection frequency. The desired RF signal f_d mixes with the LO signal to produce an IF frequency. The undesired signal's second harmonic, $2f_u$, called the half-IF spur, mixes with the LO signal's second harmonic to produce an IF frequency, $2LO-2f_u$, shown by the dotted spectral line. This half-IF frequency interferes with the reception of the desired RF signal f_d. The interfering signal, $2f_u$, is also attenuated by the RF filter prior to reaching the mixer.

Because the spur is a fourth-order response, the mixer's conversion efficiency is not as good as for the desired frequency conversion. Thus, some natural protection against the half-IF spur is provided. Proper IF frequency selection is crucial to determine if filtering is possible. Selecting a low IF frequency moves the spur frequency close to the RF frequency making filtering difficult, if not impossible.

Figure 47. Half-IF Spur

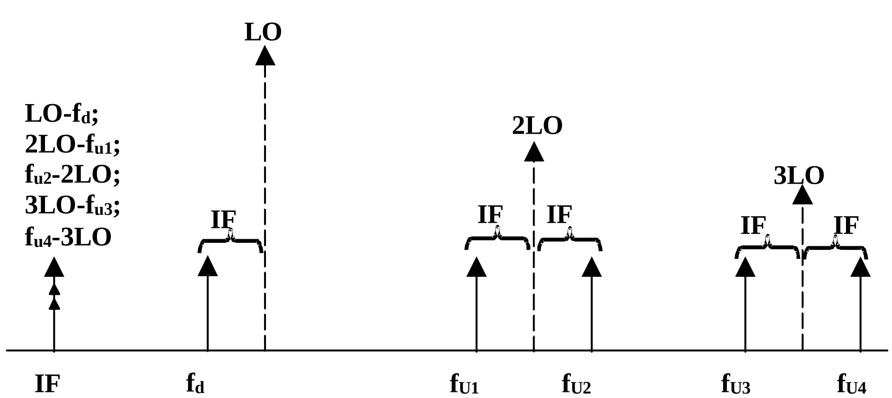


LO Spurious Responses

Additional unwanted IF responses result from mixing the LO's harmonics with RF signals spaced one IF away. The interfering RF signals that produce the unwanted IF responses are known as LO spurious responses. Figure 48 shows the frequency spectrum for the RF and LO frequencies. The LO signal and its harmonics are injected into the mixer LO port shown in dashed spectral lines. (The mixer also internally generates LO harmonics.) The desired signal f_d and all undesired signals f_{U1}, f_{U2}, f_{U3}… are available at the RF input port and shown in solid spectral lines.

Each signal at the RF port is shown to have equal signal level, thus have an opportunity for down-conversion to the IF frequency. The undesired responses are spaced an IF frequency away above and below each LO harmonic. Each mixes down directly to an IF frequency to interfere with the desired signal. LO and RF filters protect from LO spurs. Proper selection of the IF also determines whether or not filtering in the RF path is effective against LO spurs.

Figure 48. Frequency Spectrum for RF and LO Frequencies



Self-Quieters

Self-quieters, heterodynes, and birdies are names that describe internally generated interference. The receiver may not be able to receive weak signals because the receiver has locked on to some internally generated signal. Common sources of self-quieters include microcomputer clocks with square wave signals having substantial power in higher order harmonics. Some of the harmonics fall within the RF passband and interfere with the desired signal. As higher clock frequencies are introduced into communications equipment, lower order harmonics with more power fall into the RF passband.

Analog audio communications equipment produces audible noise when no desired RF carrier is present. For example, random noise generated by the source termination and by circuits in an FM receiver has random frequency components. The FM demodulator produces random voltages across the speaker proportional to the random frequency components—noise is heard. If an unmodulated carrier is introduced at the receiver's input, the audible noise drops in volume. The receiver has been quieted. With the unmodulated carrier at the input to the FM receiver, a DC voltage is produced across the speaker, thus producing no sound (quieted). If a clock harmonic falls within the RF passband, the radio is again quieted, but by an internally generated signal. If an external unmodulated desired RF signal is again introduced at the receiver input, it beats against the clock harmonic, thus generating an audible tone known as a birdie.

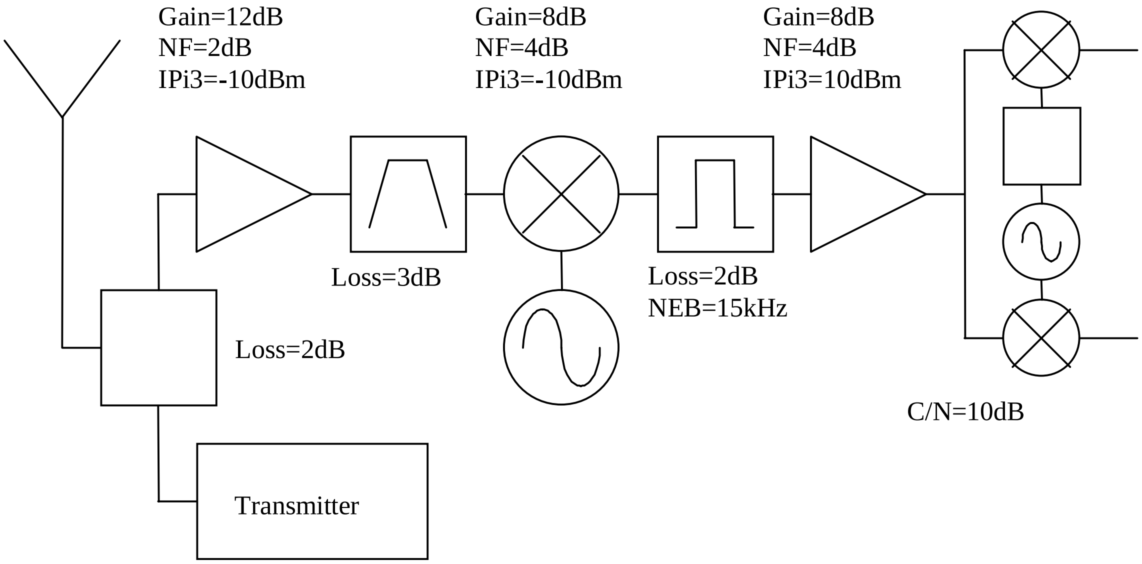
The mechanism is the same for digital receivers as for analog, but the only evidence of the self-quieter is an increased bit error rate on the self-quieted channels.

To protect against self-quieters, a clock shifter is provided to shift the clock harmonic out of the RF passband for specific channels. The channels requiring a clock shift are programmed into a look-up table for computer controlled systems. Rounding off the clock's square edges reduces power contained in higher order harmonics, thus providing some protection.

Receiver Design Trade-Offs

This section compares sensitivity and spur-free dynamic range when the RF filter follows the RF amplifier and when the RF filter precedes the RF amplifier. Figure 49 shows an example of a receiver block diagram. The RF filter protects the mixer from image and half-IF spurs. Sensitivity and spur-free dynamic range are compared when the filter follows the RF amplifier and when the RF filter precedes the RF amplifier. The block diagram shows the filter following the RF amplifier.

Figure 49. Receiver Block Diagram

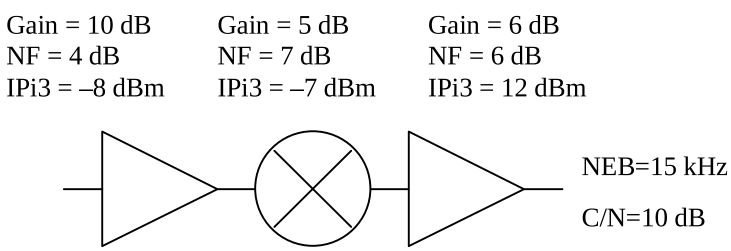


First, the total gain for the system G_{sys} is calculated by simply adding all the gains of each stage together.

$$G_{sys} = -2 + 12 - 3 + 8 - 2 + 8 = 21\,\mathrm{dB} \tag{97}$$

Next, the losses in front of gain stages are combined with those gain stages. The six-stage network combines into a three-stage network, as shown in Figure 50. A loss in front of a gain stage decreases gain dB for dB and increases noise figure and input intercept point dB for dB (see Figure 50). This is true for noise figure as long as additional noise is not added from external sources in the lossy stage. It is true for intercept point if the loss does not generate distortion products of its own (that is, it is linear.)

Figure 50. Six-Stage Network Combined Into Three-Stage Network



Now, all dB values of gain, noise figure, and input intercept point are converted to numeric using numeric equations. (The last stage gain does not need to be converted to numeric for the cascaded noise figure or cascaded intermodulation distortion equations.) The system numeric gains are given by:

$g_1 = 10 \quad g_2 = 3.16$

The noise factor of each stage is:

$F_1 = 2.5 \quad F_2 = 5 \quad F_3 = 4$

Finally, the numeric input intercept points of each stage are given by:

$ip_{i1} = .158 \quad ip_{i2} = .2 \quad ip_{i3} = 15.85$

System noise factor is given by equation (51). Substituting the previous numeric values into equation (51) yields total system noise factor.

(98) $$F_T = 2.5 + \frac{5-1}{10} + \frac{4-1}{10(3.16)} = 3$$

Converting to noise figure using equation (54) yields:

(99) $$NF_T = 10\log(3) = 4.77\text{dB}$$

Third-order input intercept point is calculated using equations (87) and (88).

(100) $$\frac{1}{ip_{iT}} = \frac{1}{.158} + \frac{10}{.2} + \frac{10(3.16)}{15.85} = 58.32 \quad ip_{iT} = .017$$

(101) $$IP_{3i} = 10\log(.017) = -17.7\text{dBm}$$

The sensitivity is calculated from equation (75).

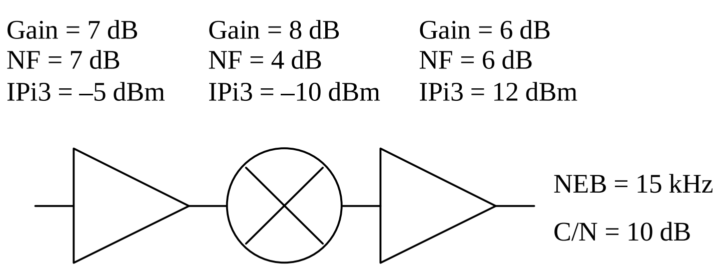
(102) $$\text{Sensitivity}_{\text{dBm}} = -174\text{ dBm} + 41.77\text{ dB} + 4.77\text{ dB} + 10\text{ dB} = -117.46\text{ dBm}$$

Spur-free dynamic range is calculated from equation (94).

(103) $$\text{SFDR} = \frac{2}{3}\left[3.3\text{ dBm} + 127.46\text{ dB} - 21\text{dB}\right] = 73\text{ dB}$$

If the filter is moved in front of the RF amplifier, the new combined receiver is shown in Figure 51.

Figure 51. New Combined Receiver



The new sensitivity and spur-free dynamic range are –114.93 dBm and 71.67 dB. Both sensitivity and SFDR where degraded by moving the RF filter.

If the modulation scheme were changed to reduce the bandwidth from 15 kHz down to 10 kHz, the sensitivity level would improve by 1.77 dB.

Increasing the RF amplifier gain of the first receiver from 12 dB to 15 dB changes sensitivity to –127.8 dBm and the SFDR to 71.43 dB.

Summary

Several parameters limit the reception of signals by the receiver. Design trade-offs are not trivial and the RF environment dictates the optimal solution. A thorough understanding of noise sources and methods of minimizing degradation allows an optimal design for signals with small amplitudes (far away). Undesired signals hinder the reception of desired signals because of circuit nonlinearity. A full understanding of nonlinearity and undesired response mechanisms is critical when providing protection against these undesired responses. A low-noise system design typically does not produce the best linearity, and high linearity typically produces more noise. Oscillator phase noise generated by the receiver oscillators also hinders reception of a desired signal, if an undesired signal with large amplitude is nearby in frequency. A thorough understanding of the receiver RF environment allows proper design specifications for optimal noise and linearity considerations.

Glossary

The following terms are often used in discussions on communications receiver designs.

ACP	adjacent channel power
ADC	analog-to-digital converter
AFC	automatic frequency control
AGC	automatic gain control
AMPS	Advanced Mobile Phone Service
ASK	amplitude shift keying
AWGN	additive white Gaussian noise
Baseband	range of frequencies covered by information signal (low frequency)
BER	bit error rate
BFO	beat frequency oscillator
BPSK	binary phase shift keying
BSC	base-station controller
CAD	computer-aided design
CAE	computer-aided engineering
CDMA	code-division multiple access
CDPD	cellular digital packet data
C/I	carrier-to-interference ratio
C/N	carrier-to-noise ratio
CSMA	carrier sense multiple access

DAC	digital-to-analog converter
DECT	Digital European Cordless Telecommunications
DDS	direct-digital synthesizer
DPSK	differential phase-shift keying
DQPSK	differential quadrature phase shift keying, a modulation format in which QPSK signals are differentially shifted by 45 degrees.
DSP	digital signal processing
DSSS	direct-sequence spread spectrum
E_b	effective bit energy
EIRP	effective isotropic radiated power
EMC	electromagnetic compatibility
EMI	electromagnetic interference
ERP	effective radiated power
F	noise factor
FDD	frequency-division duplex
FDM	frequency division multiplex
FEC	forward error correction
FHSS	frequency-hopping spread spectrum
FSK	frequency-shift keying
GFSK	Gaussian frequency-shift keying
GMSK	Gaussian minimum-shift keying
GPS	global positioning system
GSM	Global System for Mobile Communications
HPA	high-power amplifier
HTS	high-temperature superconductor
IF	intermediate frequency
IM	intermodulation
IMD	intermodulation distortion
I/Q	in-phase/quadrature
IR	infrared
IS	Interim Standard
IS-54	Interim Standard 54 for dual-mode TDMA/AMPS
IS-95	Interim Standard 95 for dual-mode CDMA/AMPS
ISI	intersymbol interference
JDC	Japanese Digital Cellular

LHCP	left-hand circular polarization
LMDS	local multichannel distribution system
LMR	land mobile radio
LNA	low-noise amplifier
LO	Local oscillator
LOS	line of sight
MAC	medium-access control
MAP	mobile application part
MDS	multipoint distribution system
MFLOPS	million floating-point operations per second
MIPS	million instructions per second
MMDS	multichannel multipoint distribution system
MPSK	minimum phase-shift keying
MSK	minimum-shift keying
MUX	multiplexer
NADC	North American Digital Cellular
NAMPS	narrowband Advanced Mobile Phone Service
NF	noise figure
NMT	Nordic Mobile Telephone
NRZ	non-return to zero
OCXO	oven-compensated crystal oscillator
OOK	on-off keying
OQPSK	offset quadrature phase-shift keying
PACS	personal advanced communications systems
PAM	pulse amplitude modulation
PBX	private branch exchange
PCM	pulse code modulation
PCN	Personal Communications Network
PCS	Personal Communications Services
PDF	probability density function
PEP	peak envelope power
PHP	Personal HandyPhone
PHS	Personal HandyPhone System
PLL	phase lock loop
PMR	professional mobile radio
PN	pseudorandom noise
PPB	parts per billion
PPM	parts per million
PSK	phase-shift keying

PWM	pulse-width modulation
Q	quality factor
QAM	quadrature amplitude modulation
QPSK	quadrature phase-shift keying
RHCP	right-hand circular polarization
RMS	root mean square
RSSI	received signal-strength indicator
SAW	surface-acoustic wave
SDMA	space-division multiple access
Selectivity	receiver rejection capability to undesired signals
SINAD	signal-to-noise plus distortion
SMR	specialized mobile radio
SNR	signal-to-noise ratio
SS	spread spectrum
SSB	single sideband
TACS	Total Access Communications System
TCXO	temperature-compensated crystal oscillator
TDD	time-division duplex
TDMA	time-division multiple access
TETRA	trans-European trunked radio system
T/R	transmit/receive
TWT	traveling-wave tube
TWTA	traveling-wave tube amplifier
UHF	ultrahigh frequency
VCO	voltage-controlled oscillator
VCXO	voltage-controlled crystal oscillator
VHF	very high frequency
VPSK	variable phase-shift keying
VSAT	very-small-aperture terminal
VSB	vestigial sideband
VVA	voltage-variable attenuator
WAN	wide-area network
WLAN	wireless local-area network
WLL	wireless local loop

References

Les Besser, *Applied RF Techniques I*, Besser Associates, Los Altos, CA, 1995.

Les Besser, *Passive Component Modeling*, Besser Associates, Los Altos, CA, 1997.

Gary A. Breed, *Receiver Basics—Part 2: Fundamental Receiver Architectures*, RF Design, March 1994, pp. 84-89.

Scott R. Bullock, *Transceiver System Design for Digital Communications*, Noble Publishing Company. Tacker, GA, 1995.

Steve Ciccarelli, *Predict Receiver IM in the Presence of LO Phase Noise*, Microwaves & RF, September 1996, pp. 86-90.

Rifaat A. Dayem, *Mobile Data & Wireless LAN Technologies* (Prentice-Hall Series in Computer Networking and Distributed Systems), Prentice-Hall PTR, Upper Saddle River, NJ, 1997.

Ian Getreu, *Modeling the Bipolar Transistor*, Tektronix Inc., Beaverton, OR, 1976.

Jerry D. Gibson, Editor, *The Mobile Communications Handbook*, CRC Press, Boca Raton, FL, 1996.

Jerry D. Gibson, Editor, *The Communications Handbook*, CRC Press, Boca Raton, FL, 1997.

Alex Glatfelter, *Computing Image Filter Requirements for Digital Cellular Radios*, RF Design, February 1996, pp. 76-79.

Guillermo Gonzalez, *Microwave Transistor Amplifiers Analysis and Design*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1984.

Steve Hamilton, *Introduction to Nonlinear Device Model Extraction Techniques for Circuit Designers*, Besser Associates, Los Altos, CA, 1997.

Peter Hatzold, *Understanding Digital Modulation for Mobile-Radio Systems*, Wireless Systems Design Master Reference, July, 1997. pp. 46-54.

Paul Heleine and Nils Nazoa, *Estimate IM In Cellular Radio Applications*, Microwaves & RF, April 1997, pp. 54-60.

An Introduction to Digital Radio Principles, Practical Problems, and Measurement Solutions, Hewlett Packard AN 355A, 1992.

Digital Modulation in Communication Systems—An Introduction, Hewlett Packard AN 1298, 1997.

Fundamentals of RF and Microwave Noise Figure Measurements, Hewlett Packard AN 57-1, 1983.

Noise Figure Measurement Accuracy, Hewlett Packard AN 57-2, 1988.

Noise Figure Measurement Principles and Applications, Hewlett Packard, 1989.

H. L. Krauss, W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, John Wiley & Sons, NY, 1980.

Lawrence E. Larson, Ed., *RF and Microwave Circuit Design for Wireless Communications*, Artech House, Norwood, MA, 1996.

Stephen A. Maas, *Microwave Mixers*, 2nd Ed., Artech House, Inc., Norwood, MA, 1993.

Asha Mehrotra, *Cellular Radio Performance Engineering* (Artech House Mobile Communications Series), Artech House, Norwood, MA, 1994.

William E. Pastori, *Receiver Noise Measurements*, Ontario, Canada, Maury Microwave Corporation, Short Course on RF and Microwave Measurements for Wireless Applications, ARFTG-NIST, December 2-3, 1997.

Jeffrey Pawlan, *A Tutorial on Intermodulation Distortion and Non-Linearity in RF Systems*, RF Design, February 1996, pp. 71-74.

Ken Payne, *LNAlite Template Guide*, Hewlett Packard MDS Reference Manual, Boynton Beach, FL, 1998.

Ken Payne, *Low Power RF Amplifier Design for Paging or Cellular Telephone Applications*, Hewlett Packard Application Note, Boynton Beach, FL, 1998.

Ken Payne, *Low Power RF Mixer Design for Paging or Cellular Telephone Applications*, Hewlett Packard Application Note, Boynton Beach, FL, 1997.

Ken Payne, *RF Mixer Design Using Nonlinear Simulation*, Proceedings of the Wireless & Microwave Technology 1997 Conference, 1997.

Ken Payne, *RF Productivity Improvement*, Besser Associates, Boynton Beach, FL, 1998.

Ken Payne, *Wireless RF Component and System Measurements*, Besser Associates, Boynton Beach, FL, 1998.

Raoul Pettai, *Noise in Receiving Systems*, John Wiley & Sons, New York, 1984.

David M. Pozar, *Microwave Engineering*, 2nd Edition, John Wiley & Sons, New York, 1996.

John G. Proakis and Masoud Salehi, *Communications Systems Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1994.

Siegmund M. Redl, Matthias K. Weber, Malcolm W. Oliphant, *An Introduction to GSM*, Artech House, Norwood, MA, 1995.

Theodore S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall PTR, Upper Saddle River, NJ, 1996.

Ulrich L. Rohde, Jerry Whitaker, and T. T. N. Bucher, *Communications Receivers*, 2nd Edition, McGraw-Hill, New York, 1997.

William E. Sabin and Edgar O Schoenike, Editors, *Single Sideband Systems & Circuits*, McGraw-Hill, New York, 1995.

R. C. Sagers, *Intercept Point and Undesired Responses*, 32^{nd} IEEE Vehicular Technology Conference, May 23-25, 1982.

M. R. Spiegel, *Mathematical Handbook of Formulas and Tables*, Schaum's Outline Series, McGraw-Hill Book Company, NY 1968.

G. B. Thomas Jr. and R. L. Finney, *Calculus and Analytic Geometry*, 5th Ed., Addison-Wesley Publishing Company, Menlo Park, CA, 1980.

James Tsui, *Digital Techniques for Wideband Receivers*, Artech House, Norwood, MA, 1995.

Peter Vizmuller, *RF Design Guide: Systems, Circuits, and Equations*, Artech House, Norwood, MA, 1995.

Steve Winder, *Single Tone Intermodulation Testing*, RF Design, December 1993, pp. 34-44.

M. E. Vanvalkenburg, *Network Analysis*, 3rd Ed., Prentice-Hall, Inc., Englewood Cliffs, NJ, 1974.

IMPORTANT NOTICE

Texas Instruments and its subsidiaries (TI) reserve the right to make changes to their products or to discontinue any product or service without notice, and advise customers to obtain the latest version of relevant information to verify, before placing orders, that information being relied on is current and complete. All products are sold subject to the terms and conditions of sale supplied at the time of order acknowledgement, including those pertaining to warranty, patent infringement, and limitation of liability.

TI warrants performance of its semiconductor products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are utilized to the extent TI deems necessary to support this warranty. Specific testing of all parameters of each device is not necessarily performed, except those mandated by government requirements.

CERTAIN APPLICATIONS USING SEMICONDUCTOR PRODUCTS MAY INVOLVE POTENTIAL RISKS OF DEATH, PERSONAL INJURY, OR SEVERE PROPERTY OR ENVIRONMENTAL DAMAGE ("CRITICAL APPLICATIONS"). TI SEMICONDUCTOR PRODUCTS ARE NOT DESIGNED, AUTHORIZED, OR WARRANTED TO BE SUITABLE FOR USE IN LIFE-SUPPORT DEVICES OR SYSTEMS OR OTHER CRITICAL APPLICATIONS. INCLUSION OF TI PRODUCTS IN SUCH APPLICATIONS IS UNDERSTOOD TO BE FULLY AT THE CUSTOMER'S RISK.

In order to minimize risks associated with the customer's applications, adequate design and operating safeguards must be provided by the customer to minimize inherent or procedural hazards.

TI assumes no liability for applications assistance or customer product design. TI does not warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right of TI covering or relating to any combination, machine, or process in which such semiconductor products or services might be or are used. TI's publication of information regarding any third party's products or services does not constitute TI's approval, warranty or endorsement thereof.